



## Knowledge of knots: shapes in action

Roberto Casati

### ► To cite this version:

Roberto Casati. Knowledge of knots: shapes in action. Kutz, O., Bhatt, M., Borgo, S., Santos, P. Shapes 2.0: The Shapes of Things / Universal Logic, Apr 2013, Rio de Janeiro, Brazil. Universal Logic, pp.3-20, 2013. <ijn\_00911999>

**HAL Id: ijn\_00911999**

**[http://jeannicod.ccsd.cnrs.fr/ijn\\_00911999](http://jeannicod.ccsd.cnrs.fr/ijn_00911999)**

Submitted on 1 Dec 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **SHAPES 2.0**

## The Shape of Things

*Workshop held at the*  
World Congress and School on Universal Logic  
April 3-4 2013

Rio de Janeiro, Brazil

### **Editors**

Oliver Kutz      Mehul Bhatt      Stefano Borgo      Paulo Santos





## PREFACE

### Shape, Form, and Structure

Shape, Form, and Structure are some of the most elusive notions within diverse disciplines ranging from humanities (literature, arts) to sciences (chemistry, biology, physics etc.) and within these from the formal (like mathematics) to the empirical disciplines (such as engineering and cognitive science). Even within domains such as computer science and artificial intelligence, these notions are replete with commonsense meanings (think of everyday perception and communication), and formalisations of the semantics and reasoning about shape, form, and structure are often *ad hoc*. Whereas several approaches have been proposed within the aforementioned disciplines to study the notions of shape, form and structure from different standpoints, a comprehensive formal treatment of these notions is currently lacking and no real interdisciplinary perspective has been put forward.

The workshop series SHAPES provides an interdisciplinary platform for the discussion of all topics connected to shape (broadly understood): perspectives from psycho-linguistics, ontology, computer science, mathematics, aesthetics, and cognitive science, amongst others, are welcome to contribute and participate in the workshops. We seek to facilitate a discussion between researchers from all disciplines interested in representing shape and reasoning about it. This includes formal, cognitive, linguistic, engineering and/or philosophical aspects of space, as well as their application in the sciences and in the arts.

We also welcome contributions on the relationship of shape representations at different levels of detail (e.g. 2D, 3D) and in different logics, and with respect to different qualitative and quantitative dimensions, such as topology, distance, symmetry, orientation, etc.

### Form and Function in Natural and Artificial Systems

Within the philosophy and practice of design, the ontological notions of shape, form and structure have a further role of constraining function, malfunction, and behaviour of things. In this perspective, the decision-making process in design is a trade-off between physical, logical and cognitive laws and constraints that intertwine shapes and functionalities. Here, the spatio-linguistic, conceptual, formal, and computational modeling of shape serves as a crucial step towards the realization of functional affordances. This line of thought extends to several other disciplines concerned not only with the design of technical systems, but also with the understanding of biological as well as socio-technical systems. For instance, in biochemistry the shape of molecular entities (proteins, small molecules) has a direct effect on their interactions which give rise to the capacities they can manifest and, in turn, to the processes of life and death. Representing and reasoning about the shapes and realizable functionalities of these entities is essential to understand basic biological processes. Of special importance, in this as well as other

contexts, is the understanding of shape complementarity, that is, categorising the shapes of holes as well as the shapes of the entities that can fit into those holes, which can either facilitate or block the functionality of the overall system.

The results of this workshop will stimulate and facilitate an active exchange on interdisciplinary applications, ideas, approaches, and methods in the area of modelling shape, form, pattern and function. The format of the workshop combined invited speakers, peer-reviewed full contributions, as well as short position and demo papers, and allowed ample time for open discussions amongst the participants. Topics covered included:

**Linguistics / Philosophy** shape and form in natural language; differences between shape, form, structure, and pattern; shape in natural and artificial objects.

**Cognition** shape perception and mental representation; gestalt vs. structuralist understanding of shape cognition; perception and shape (e.g. identifying objects from incomplete visual information); affordances, dispositions, and shape.

**Logics, Spatial Representations** formal characterisations of shape and form; logics for shape: e.g. fuzzy, modal, intensional; logics for topology, symmetry, shape similarity; design semantics, spatial semantics; shape and 3D space; shape and space in cognitive assistance systems.

**Ontology** ontologies and classifications of shapes; ontological relations among shape, objects and functions; patterns as shapes of processes; forms and patterns in ontology.

**Applications** *Biology & Chemistry*: molecular shapes, shape in anatomy and phenotype definitions, shape complementarity between objects and holes, shape in medical image analysis and annotation.

*Visual Art and Aesthetics*: shape in Film and Photography; shape in computational creativity.

*Naive Physics and Geography*: e.g. qualitative classifications of shapes of geographic objects.

*Design & Architecture*: shape grammars; CAD, symmetry and beauty in architectural design.

*Engineering*: formal shape analysis in engineering processes.

The workshop SHAPES 2.0 followed a successful first event held at CONTEXT 2011 in Karlsruhe, Germany.<sup>1</sup>

SHAPES 2.0 grew significantly in its second installment<sup>2</sup>, running as a full two-day workshop, and attracting a total of 23 contributed submissions of which we selected 14 for presentation at the workshop, with an additional 5 invited contributions. We thank all the speakers for their great presentations, and the audience for generating very lively and fruitful discussions.

---

<sup>1</sup>See <http://cindy.informatik.uni-bremen.de/cosy/events/shapes/> for the workshop website. The proceedings have been published as Vol. 812 of the CEUR workshop proceedings, edited by Janna Hastings, Oliver Kutz, Mehul Bhatt, and Stefano Borgo, see <http://ceur-ws.org/Vol-812/>.

<sup>2</sup>See <http://cindy.informatik.uni-bremen.de/cosy/events/shapes2/> for the workshop website

## Acknowledgements

We would like to thank the program committee members and the additional reviewers for their timely reviewing. We thank our invited keynote speakers—Roberto Casati, Roberto M. Cesar Jr, Simon Colton, Antony Galton, and Barbara Tversky—for their support and contributions. We also thank the UniLog conference, in particular Jean-Yves Béziau, for hosting the second edition of SHAPES, and for generously providing free conference registration to our keynote speakers and workshop organisers.

SHAPES 2.0 has been generously sponsored by the following organisations:

- CAPES – Coordination for the improvement of Higher-level Personnel ([www.capes.gov.br](http://www.capes.gov.br))
- CNPq – National Council for Scientific and Technological Development ([www.cnpq.br](http://www.cnpq.br))
- International Association for Ontology and its Applications (IAOA) ([www.iaoa.org](http://www.iaoa.org))

SHAPES is an initiative of the IAOA Special Interest Group:  
Design Semantics ([www.designsemantics.org](http://www.designsemantics.org))

Mehul Bhatt and Oliver Kutz acknowledge support of the German Research Foundation (DFG) via the Spatial Cognition Research Center (SFB/TR 8) co-located at the University of Bremen, and the University of Freiburg, Germany (<http://www.sfbtr8.spatial-cognition.de>).

## Shapes 2.0 – Organisation

### Programme Chairs

Oliver Kutz	Spatial Cognition Research Center (SFB/TR 8) University of Bremen, Germany
Mehul Bhatt	Cognitive Systems (CoSy), and Spatial Cognition Research Center (SFB/TR 8) University of Bremen, Germany
Stefano Borgo	Laboratory for Applied Ontology (LOA), ISTC-CNR, Trento, Italy
Paulo Santos	Artificial Intelligence in Automation group, Centro Universitário da FEI Sao Paulo, Brazil

### Programme Committee

Colin Batchelor	Royal Society of Chemistry, Cambridge, UK
John Bateman	University of Bremen, Germany
Thomas Bittner	SUNY, Buffalo, USA
Walter Carnielli	State University of Campinas, Brazil
Hannah Dee	Aberystwyth University, UK
Christian Freksa	University of Bremen, Germany
George Gkoutos	University of Cambridge, UK
Michael Grüninger	University of Toronto, Canada
Giancarlo Guizzardi	University of Espírito Santo, Brazil
Janna Hastings	University of Geneva, Switzerland
Shyamanta Hazarika	Tezpur University, India
Roberto Hirata Jr.	University of Sao Paulo, Brazil
Ingvar Johansson	Umea University, Sweden
Joao Kogler	University of Sao Paulo, Brazil
Kai-Uwe Kühnberger	University of Osnabrück, Germany
Frieder Nake	University of Bremen, Germany
Susanna Siegel	Harvard University, USA
Achille Varzi	Columbia University, USA
Pieter Vermaas	Delft University of Technology, The Netherlands

### Additional Reviewers

Pawel Garbacz	The John Paul II Catholic University of Lublin, Poland
Torsten Hahmann	University of Toronto, Canada
Valquiria F. Pereira	University of Sao Paulo, Brazil
Vinicius Romanini	University of Sao Paulo, Brazil
Flávio Soares Corrêa Da Silva	University of Sao Paulo, Brazil
Marcio Lobo Netto	University of Sao Paulo, Brazil
Nicolau Leal Werneck	Geekie Software, Brazil
Carlos da Silva dos Santos	Universidade Federal do ABC, Brazil
Yossi Zana	Universidade Federal do ABC, Brazil





# CONTENTS

<b>Invited Speakers</b>	I
Roberto Casati: <b>Knowledge of knots: shapes in action</b>	3
Roberto Marcondes Cesar Jr.: <b>Graph-based pattern recognition and applications</b>	21
Simon Colton: <b>On Shape Poems and the Shape of Poems: A Computational Creativity Perspective</b>	23
Antony Galton: <b>Prolegomena to an Ontology of Shape</b>	29
Barbara Tversky: <b>Lines, Shapes, and Meaning</b>	41
<b>Contributed Papers</b>	47
Ulf Krumnack, Angela Schwering, Kai-Uwe Kühnberger, Helmar Gust, Ahmed Abdel-Fattah, Tarek Besold, Martin Schmidt, and Stefan Schneider: <b>Sketch Learning by Analogy</b>	49
Mehul Bhatt and Carl Schultz: <b>The Shape of Empty Space: Human-centred cognitive foundations in computing for spatial design</b>	59
Ophelia Deroy and Malika Auvray: <b>A new Molyneux's problem: Sounds, shapes and arbitrary crossmodal correspondences</b>	61
Sandro Rama Fiorini, Mara Abel, and Peter Gärdenfors: <b>Structure, Similarity and Spaces</b>	71
Nicola Guarino: <b>Local Qualities, Quality Fields, and Quality Patterns: A Preliminary Investigation</b>	75
Janna Hastings, Colin Batchelor, and Mitsuhiro Okada: <b>Shape Perception in Chemistry</b>	83
Silvia Likavec: <b>Shapes as property restrictions and property-based similarity</b>	95
Alessandra Melas: <b>The Shape of Absolute Coincidences. Salmon's Interactive Fork Model as shape of coincidental processes</b>	107
José Gustavo Morales and Matías Saracho: <b>The Role of Shape in Problem-Solving Activities in Mathematics</b>	117
Robert J. Rovetto: <b>Shaping up: The Phenotypic Quality Ontology and Cross Sections</b>	125
Stefan Schneider, Ulf Krumnack, and Kai-Uwe Kühnberger: <b>Dynamic Assembly of Figures in Visuospatial Reasoning</b>	137
Carl Schultz and Mehul Bhatt: <b>Declarative Computing with Shapes and Shadows</b>	143
Eric Tremault: <b>Mental "structures" in the Berlin school of Gestalt Psychology: can sensation be described as "structural"?</b>	151
Christoph Zetsche, Konrad Gadzicki and Tobias Kluth: <b>Statistical Invariants of Spatial Form: From Local AND to Numerosity</b>	163





## Invited Keynotes



# Knowledge of knots: shapes in action

Roberto Casati\*

*Institut Jean Nicod (CNRS-EHESS-ENS)*

Keywords: Knots, shapes, topology, processes, action

**Abstract:** Logic is to natural language what knot theory is to natural, everyday knots. Logic is concerned with some cognitive performances; in particular, some natural language inferences are captured by various types of calculi (propositional, predicate, modal, deontic, quantum, probabilistic, etc.), which in turn may generate inferences that are arguably beyond natural logic abilities, or non-well synchronized therewith (eg. *ex falso quodlibet*, material implication). Mathematical knot theory accounts for some abilities - such as recognizing sameness or differences of some knots, and in turn generates a formalism for distinctions that common sense is blind to. Logic has proven useful in linguistics and in accounting for some aspects of reasoning, but which knotting performances are there, over and beyond some intuitive discriminating abilities, that may require extensions or restrictions of the normative calculus of knots? Are they amenable to mathematical treatment? And what role is played in the game by mental representations? I shall draw from a corpus of techniques and practices to show to what extent compositionality, lexical and normative elements are present in natural knots, with the prospect of formally exploring an area of human competence that interfaces thought, perception and action in a complex fabric.

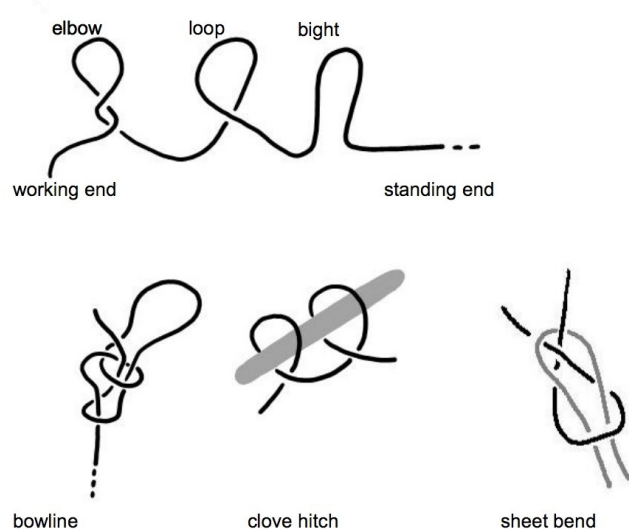


Fig 1. Some of the items we are going to discuss in this paper, listed here to assist the reader.

\*Roberto Casati, Institut Nicod CNRS-EHESS-ENS, Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris, France. Phone: +33 144322695, E mail: casati@chess.fr

The shoelace knot is the most common mildly complex knot everyone learns to tie. Most of us can tie it with closed eyes; our fingers somewhat know what to do.

I discovered recently that it can be tied in at least two more ways, over and above what I thought to be the canonical way. My youngest daughter was taught in school that one can tie a Plain Knot on two bights (“bunny ears”), and I found a number of tutorials on the web showing how to create half loops on your fingers to tie the knot in a single, swift movement – something I learned to do myself with huge intellectual pleasure.

I also decided as of late that I was able to analyse the shoelace knot. It is, actually, a composed knot: a *Plain Knot* followed by a *running Half Hitch* on a *bight*. These are semitechnical notions that I learned from sailing practice. More technically, bordering topological notions, thus cautiously, we can say that the shoelace knot decomposes into here a sequence of a “genuine” knot and an “unknot”.

Knowing how it decomposes made me a bit ambitious. Can the shoelace knot be improved upon? I somewhat succeeded in getting rid of the Plain Knot (which I dislike, like many sailors) and ensuring some stability by tying two Half Hitches on the bight (the latter one is once more a *running Half Hitch*, which provides easy unfastening).

The improvement is intellectually pleasant. Although we should handle mathematical notions with care, I'd say that I managed to replace a hybrid of a knot plus an unknot with something that is purely an unknot.

I think most of us appreciate that there are at least two action *atoms* in tying shoelaces. One could produce the Plain Knot without tying the running Half Hitch, and conversely. One may even understand something more – even if, I surmise, very few have ever tried this: tying first the Half Hitch, then the Plain Knot; i.e., execute the two steps of the shoelace knot in reverse order. Now, I predict that you will be surprised by the outcome: pulling the two ends, you end up with the Plain Knot! Exactly as it would happen when you pull the two ends of the shoe's knot. In both cases, the unknot disappears, and the knot stays.

The lesson from this simple example is that even if you have some understanding of the compositional structure of an action such as tying the shoe knot, you

do not thereby have an access to the end result of just any knotting procedure that involves the elements of the composition. The consequences of the atomic actions you perform are not easy to predict; not even for experts.

My purpose here is to trace the perimeter of a small research program. There are many knotting performances that one might want to explain. People tie knots, even complex ones, learn to tie knots, talk about knots, draw them, understand knot diagrams, teach knots, at various levels of expertise and conscious understanding. There is a rich set of explananda. Moreover, the examples above suggest that some decomposition, some structure is available to knotters and guides their action. The main research question is thus: what is the structure of the underlying competence that accounts for these performances?

### Knots in topology

As we search for structure, we note that the theoretical landscape is not empty. Knots are topologically interesting objects and a mathematical theory of knots has developed, providing descriptive and inferential tools to solve a number of problems. For our purposes, the main aspects of the topological account of knots are the distinction between knots and the unknot (Fig. 2), and the study of knot equivalencies. A further aspect of less concern is the peculiar classification of knots that is delivered by topology.

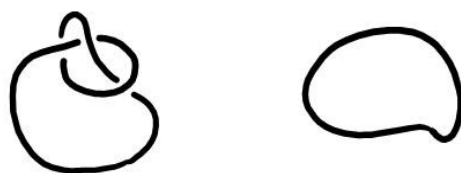


Fig. 2. Topological knot and topological unknot.

What use can be made of topological knot theory? Does it provide the appropriate framework for capturing the structure of competence? In topology, a knot is an equivalence class of knot representations, which are the planar projections (the shadows) of closed loops. The same knot can have countless representations; even when one reduces drastically

the complexity of the representation (e.g. by limiting to a small integer  $<10$  the number of crossings in the representation), the problem of identifying the members of each class has proven difficult and to date there is no algorithm that delivers a satisfactory classification [19].

Camps and Uriagereka [6] and Balari [3] have linked the complexity in knotting to that of syntax. They suggested that evidence about early human knotting practices are indirect evidence for early language use, thus proposing a link between knotting and linguistic performance. If the same computational power is assumed to underlie both knotting abilities and natural language, then evidence about knotting practice in the archeological record is evidence for at least the presence of the computational power for natural language in the brains of those who left that record. A crucial point in the argument is the recourse to topological knot theory to sustain the claim that knotting abilities require the computational power necessary for grammars, or that they share computational resources with language. We may suspend our judgement on the goal of using evidence about knotting as evidence about language (see the critiques by Lobina [15]; discussion in Balari et al [4], replies in Lobina and Brenchley [16]). At the same time we are still interested in explaining knotting competence. I would just be methodologically flexible as to the theoretical instrument we should employ. Indeed, there is no clear reason for thinking that the underlying competence is best captured by topological knot theory. There are both a generic and a specific reason for skepticism. The generic reason is purely cautionary: We have a long list of formalisms that somewhat mimic cognitive performances but in the end turn out to be quite independent from the latter and not good models thereof. Logical systems are both under and overshooting relative to people's inferential abilities. Queue theory models ideal queueing and not people's behaviour. Real-life buyers and sellers are not very well framed by rational choice theory. Coming to the point, topological classifications are misaligned with commonsense classifications [7].

Even closer to the point, the specific reason is that topological knot theory is concerned with knot equivalencies, where knots are defined over close loops in 3d-space. Ecological knots, on the other hand, are the result of transformations that take you from a situation in which there is no knot, to a

situation in which there is a knot. You do not tie topological knots, because you cannot.

An intuitive demonstration of the gap between ecological knots and topological knots is at hand. You can take a close loop and tie a Plain Knot on it (Fig. 3)



Fig. 3. A closed loop, and a knot tied on it. Not a knot for topology.

Even more dramatic are examples from real practice, for instance the cases of the Half Hitch and of the Clove Hitch. These are two most used knots. The Half Hitch, the quickest way to fix a piece of rope around an object, is fundamental in tapestry and knitting.<sup>1</sup> The Clove Hitch is a basic knot in sailing and farming. The fact is, neither is, topologically speaking, a knot.<sup>2</sup> If we resort to the graphical convention of topology, we can represent them (Fig. 4) as trivial twists in a closed loop.



Fig. 4. Half Hitch (left) and Clove Hitch (middle, right) are invisible to topology.

The Clove Hitch's advantage is precisely in the fact that it is *not* a topological knot. This means that you can tie it in the middle of a piece of rope, without having to manipulate the ends of the rope (Fig. 5).

<sup>1</sup> Tutorial received on Sep 18, 2011, at the Gobelins Tapestry in Paris.

<sup>2</sup> A word of caution. If you tie a Half Hitch at the end of a loop, you end up with a Plain Knot (a trefoil knot, which is different from an unknot). Textbooks of knotting practice tend not to distinguish between "pure" Half Hitches and Plain Knots. But this is a terminological issue. The important thing is that a Half Hitch is part of a Plain Knot.

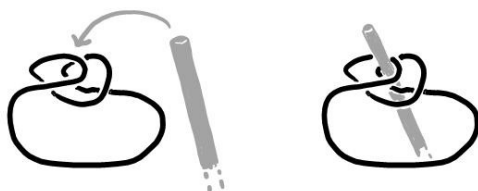


Fig. 5. Tying a Clove Hitch on a pole.

Two further aspects are that you cannot use hitches as stop knots and that you need an object around which tying them. As we shall see, there is an important notion of dependency at work here.

Our problem is thus pretty straightforward. If we find a sense for the claim that the drawing on the right hand side of Fig. 5 represents a knot, and if we accept that this knot is invisible to topology, then we need to find an alternative, non topological (or not only topological) account for our intuition that a knot is represented here.

The point of contact between topological and ecological knots concerns a small subset of explananda that have to do with recognition and categorization. A very specific ecological task is that of checking if a certain knot is the correct one (did the shipboy execute the Bowline correctly?) In this case one categorizes and assesses an equivalence, much in the same way in which topologists categorize and assess equivalences of topological knots. But, as I mentioned, this is but one of the tasks to be explained.

The Camps and Uriagereka paper [2006] makes an interesting, not uncommon assumption about the performance. It describes the execution of the Plain Knot in a way that mimicks the way the knot diagram is drawn, not the way the knot is normally tied. What is the difference? The “drawing” style consists in taking the working end of the rope and make it travel about as if it was the engine of a moving train. In real life, on the other hand, one creates a small bight, and retrieves the working end. This should interest us. What descriptions of knotting practice are to be used as good explananda? I concede that in the initial phases of learning one may use the moving train metaphor. But after a while some other gestures take over.

To sum up, the central set of problems is thus twofold:

What is the performance we want to explain?

What is the structure of the underlying competence?

The first part of what follows will be devoted to looking for interesting cases of performance, such as the understanding of knot equivalencies, description of knotting practice, etc. This is an uncharted territory.<sup>1</sup> The second part of the paper proposes a framework for dealing with the explanation of the performances. Having rejected mathematical topology as a model of competence, I shall draw on the theory of Graphic Schemes [18] in order to propose a two-step approach to knowledge of knots in sensori-motor terms, and plead for a type of mental topological representations that are process-sensitive. Knots are living memories of processes, and we need some concepts to explore their structure and constituents. We are after mid-level conceptualizations: close enough to common sense, to ecological knotting, but such as to allow for formalization. If we want to look beyond the formal toolkit of topology we do not have much of a choice. We need to start from some semi-intuitive, semi-technical ideas. Knots from topological theory will henceforth be called ‘topological knots’. Ecological knots will be just knots.

But what are knots? Let’s proceed stepwise. Metaphysically, we consider knots as physical configurations of rope (be they construed as individuals, “disturbances” [14] or properties). Not all configurations of rope are knots, of course. Besides, we take knots as configurations for which a certain relation to space is essential (if you travel in a tunnel, it does not matter for you if it is knotted or not), at least insofar as it allows for movement of the knot along the rope. Knots are stable configurations of rope that are grounded on friction, but not all such stable configurations of rope are knots. For instance, rope that is wrapped around a pole may be stable and grounded on friction, but it does not constitute a knot. Some *stabilized crossing* of rope must occur, under contextual tension. On the opposite end, a large rope jam may not count as knot for natural language: it is just a large jam. Finally, knots are what we may call *active shapes*, shapes that trap some energy. The

<sup>1</sup> For this reason, I am a bit skeptical about the conclusions one may draw about cognition of knots from the results of experiments that measure certain responses of people to the perception of knots. We do not know yet what aspects of the performance are to be explained. Cf. [22, 10]

mental representation of knots would thus be that of shapes *that store an action*. We shall rely on some intuitive understanding of the notion in what follows, within the limits set by these examples.

### What are the explananda?

We start from the explananda. The following is a mix of platitudes, personal reports, and established evidence.

(1) People tie knots. This is our starting point. Knots are extremely useful artifacts. They have various functions that rely on a basic principle, preventing rope from slipping by exerting pressure on different parts of the rope and, if they involve an object, of the object they may be tied on. Crowell [11] provides an informal digest of some of the few papers in knots physics up to 2011, in particular the seminal work of Bayman on hitches [5]. In order to work, knots must be tied in such a way as to create *nips*, friction points between parts of rope. Some parts of rope should be made to pass in loops or over other parts so that nips are formed. Typically the standing end of a piece of rope supports a load, which may serve the function of assuring a tension (this is the case with hitches). The working end of a piece of rope is in general used for tying the knot.

(2) People can untie knots – and know when a knot is so jammed, it cannot be untied, as did Alexander the Great when, according to legend, he decided it was better to cut the Gordian Knot than to try and untie it. Some simple rules for untying are: Running knots are untied by pulling the working end. Non-running knots are tied by pulling a bight.

Knowledge about knotting and knowledge about unknotting are not necessarily aligned. It looks as if one will be able to untie any knot, whereas tying specific knots requires a certain amount of training. There is, of course, an asymmetry here, related to the complexity of the task. What one is normally requested to create is a *specific* knot (say a Bowline, or a Cleat Hitch, or a Sheet Bend). One is not requested to create an unorganized knotted structure (which one may easily do by simply piling a number of simple knots and pulling the working end randomly through whichever loops are formed). Knotting and unknotting appear to require different algorithms. However, tying knowledge is useful in untying a knot. I remember that I can easily untie a Bowline; I know (but we shall see that this is no

trivial knowledge) that a Bend Sheet is a Bowline. I immediately find a way to untie a Bend sheet.

(3) We have normative intuitions about knots. In Ashley's apt words, "A knot is never 'nearly right'; it is either exactly right or it is hopelessly wrong, one or the other; there is nothing in between" [2, p.18].

(4) There is an understanding of the distinction between permanent and transitory knots. In ordinary life many knots are not permanent (shoelace knots, mooring knots, knots for climbing) and must be so designed and executed that one be able to easily untie it. Other knots, such as knots for parcels, for tapestry, for fisher nets and weaving are designed to be permanent. Most natural knots that one must quickly dispose of are unknots (the Clove Hitch). Most natural knots that one should not dispose of (stop knots) are topological knots. There is an understanding of what kind of knot is suitable for different purposes, and thus of the functional properties of each knot.

(5) People show clear degrees of expertise in tying knots. This point is less trivial, but no less true. I acquired a certain expertise; before that, I admired other people's expertise. Expertise manifests itself in speed and accuracy of the performance, in recognitional/parsing abilities, in assessment of other people's performances, in style of execution, in the ability of generalizing, in "parsing" knots one has not seen before.

(5.1) People with a limited knot repertoire face a number of knotting problems that they routinely fail to solve. A classical example is the tying up of a parcel. Without knowledge of appropriate knots and techniques one will inevitably end up with a loose rope. Another example, concerning the understanding of rope properties, is the systematic kinking of water hoses when coiling them.

(5.2) The standard knot repertoire of the large majority of adults who do not have a professional or leasurable interest in knots is very small, of two-three knots, including the Plain Knot and the shoelace knot. (Personal poll, >20 individuals.) It appears that those who learn more knots are either professionals (shepherds, sailors) or people with a hobbyistic interest (e.g. fishermen).

In general, knotters take pragmatic shortcuts. They ask, What is a knot good for? In a real life scenario, instead of connecting two pieces of ropes through the Sheet Bend, that handbooks suggest as the



appropriate solution to this problem, people link two Bowlines. This is because one seldom connects two ropes, one knows how to tie a Bowline, one does not remember the Sheet Bend, and one needs to solve a problem on the spot.

(5.3) The algorithm for tying difficult, complex knots may be forgotten after a while. (Personal observation of practice.)

(5.4) Knots appear to be cross cultural. There is a large record of knotting practices for many different purposes over and above tying objects. Knots are used as marks for measuring on ropes (whence the measuring unit 'knot'). The archeological record shows probable braids in the hair of Cro-Magnon ivory heads (upper Paleolithic, -25000). Knotted carpets date back to -3000. The Inca used since -4500 and until +650 a positional number system (Quipu or Khipu, meaning 'knot' in Quechua) based on ropes and knots [23]. Different knots had different syntactic roles. Basically (but there are complications) a knot denotes a unit; series of knots represent a number between zero (no knot) to nine (nine knots). The end of a numeral was denoted by a Figure-of-Eight Knot. The value of a Many-Turn Long Knot was given by the number of turns. What matters for our purposes is the use of different knots, two of which are pretty standard (the Plain Knot and the Figure of Eight Knot). The fact that the same knots are used in different and distant cultures can be the result of cultural transmission, but more simply can be just a consequence of the fact that the space of possible solutions is not much populated at the “easy” end.

(5.5) Animals do not appear to be able to tie knots, with the remarkable exception of great apes in captivity [13]. There is some reason to suppose that this cultural habit is imported from humans. “Takanoshi Kano, a bonobo specialist, notes: “. . . I wonder wild apes may meet need to make a knot, and also you should notice that knot-like objects for apes to untie do not exist in wild situations” (p. 626-627). Herzfeld and Lestel studied the behavior of Wattana, an orangutan at the Paris Ménagerie of Jardin des Plantes. Using hands, feet and mouth Wattana tied half-Hitches, simple knots and even shoelace knots, and created some assemblages. *“Her knots were not restricted to single ones; she also made double and triple knots. Some of them were even more complex, for she passed the ends back and forth through the loops already formed. She also sometimes wrapped a string around another string held between her two*

*feet, passing the string back and forth, making loops and then passing one end of the string through one or another of the loops already made before pulling it taut. One might call this a sort of “interlacing”, a form of weaving”* (p. 631) Two facts are worth mentioning. First, Wattana used knots as projectiles (they increase locally the mass of rope). This indicates that there are practical, noncognitive attractors in knot tying that may not be part of any planning. Arguably, Wattana has made, and made use of, an interesting discovery in naïve physics. Second, Tübo, a fellow young male, untied some of Wattana's knots (p. 643). Knotting acquired social relevance.

(6) Children start tying knots at age 3-4, have a long learning phase, and a slow performance for some years (informal poll of kindergarten teachers). Strohecker [20] is a study of instruction of children in an experimental setting.

(7) Language. People teach knots by showing them but also by accompanying the ostension with a description of the algorithm that generates the knot (playing at the interface between action and the conceptual system.) It is also possible to describe the knot, i.e. the configuration of tied up rope, the structure of the knot – both in a view-dependent and in a view-independent way. Incidentally, when a manual explains a knot, it normally talks about the movement of the rope, not about the hand movement.

(8) People can see *some* knot equivalencies/differences by just visually inspecting knots. Topologists for that matter, are skilled at that. Expertise plays an important role here. In learning to solve graphical knot equivalencies, topologists make use of the Reidemeister moves (Fig. 6).

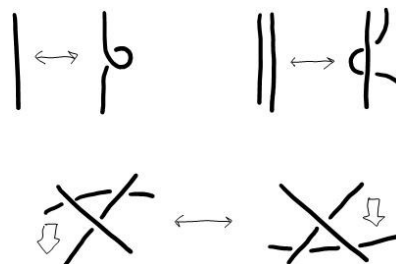


Fig. 6. The three Reidemeister moves.

The configurations linked by double arrows in Fig. 6 are local moves that do not change the corresponding topological knot and can be

interchanged in a graphic representation of a knot. Slightly more formally, the three Reidemeister moves are sufficient to connect any two diagrams that represent the same type of knot (they are “shadows” of 3D movements in the knots). The Reidemeister Theorem states that “If one knot can be transformed into another knot by continuous manipulation in space, the same result can be obtained by a manipulation whose projection consists uniquely of Reidemeister moves and trivial manipulations of the diagram in the plane” [19, p. 41]. One simple hypothesis is that after a learning phase topologists interiorize the Reidemeister moves (a discussion of topological knots in cognition in De Toffoli and Giardino [12]).

(9) People may be blind to some knot equivalencies. I want to offer two cases concerning ecological knots.

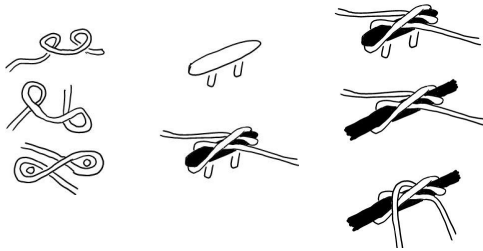


Fig. 7. The equivalence of Cleat Hitch and Clove Hitch.

The Cleat Hitch and the Clove Hitch are the same unknot (Fig. 7). The difference is in the fact that the Cleat Hitch takes advantage of the geometry of the cleat.

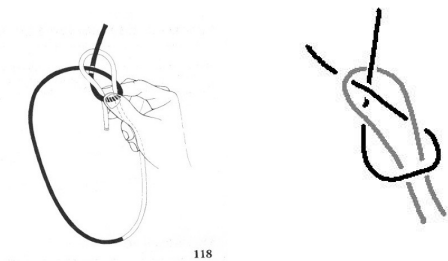


Fig. 8. The equivalence of Bowline and sheet bend. (Left figure from Asher [1], who uses two different colors for the two ends of the rope in order to point the similarity).

The Bowline and the sheet bend are the same knot (Fig. 8). The only difference is that the Bowline is tied on a single piece of rope, whereas the sheet bend is used for tying together two disconnected pieces of rope.

Although the knots involved are relatively simple, and although the equivalences have been noted in some texts ([2], [1]), knotters and many knot handbooks are largely unfamiliar with these equivalencies. I was instructed by one of my knotting teachers about the Cleat Hitch/Clove Hitch equivalence, and still find it a bit surprising. The Cleat/Clove Hitch equivalence is in a sense a purely topological equivalence; we all sense that the shape of the object has something to do with the difference. More about this later.

Some dimensions for measure could be tentatively introduced here, in reference to the population of experienced topologists. Knots can be graded according to intrinsic complexity. But they can be graded according to the subjective difficulty in parsing them as well. Thus, even the unknot (by definition, the simplest case) can be presented in ways that make it hard to parse (Fig. 9).

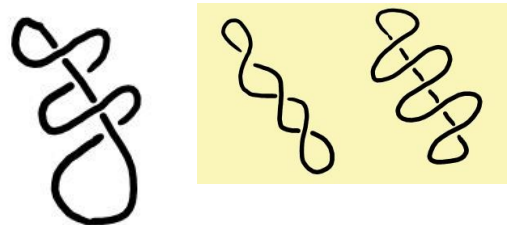


Fig. 9. Left, the unknot, under a difficult presentation. Right, two easy-to parse representations of the unknot.

Sossinsky [19] reports that only advanced algebraic techniques made it possible to show that two particular knot representations, that were considered for more than a century to belong to different knots, turned out to be in the same equivalence class.

Looking beyond knots, people have some sub-personal and personal access to topological equivalences presented visually [9] (some caveats in [8]). But people do not have access – neither personal, nor subpersonal – to relatively simple topological equivalences. Casati and Varzi [7] presented a number of cases of topologically

equivalent objects that are seen as having quite different holes in it, even pretty simple ones (Fig. 10).

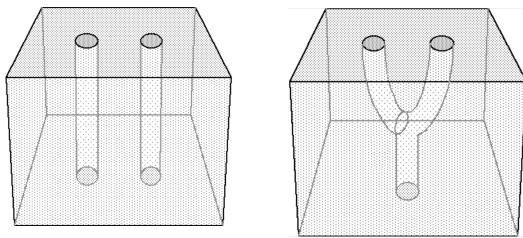


Fig. 10. The two cubic items are topologically equivalent, but they do not appear to be deformable into each other without cutting or gluing.

Topologists must train themselves to assess topological equivalencies (in particular in the case of knots, but not limited to that case) and the Reidemeister moves are meant to be an aid to pen and pencil reasoning. They so do by providing a framework to decompose any intuitive move and thus treat it mathematically.

(10) As a particular case of the previous point, expertise can be context-bound. Draftsmen who specialize on faces may be poor at drawing trees [18]. Skilled knot topologists may overlook mistakes in representations of sailors's knots. Fig P1.e of Sossinsky [19] wrongly represents a sheet bend – the “knot” will definitely untie if pulled (Fig. 11).

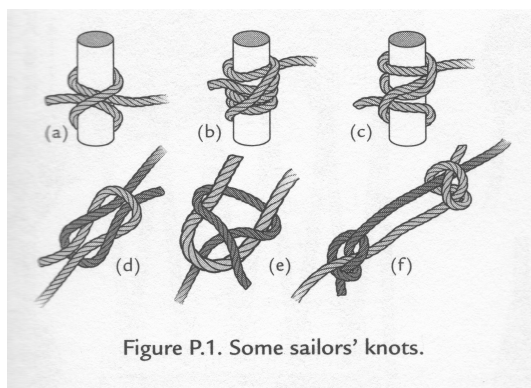


Fig. 11. Reproduced from Sossinsky [19]. The purported “sheet bend” represented in (e) will not hold.

(11) The same knot can be tied in different ways.

There is the train-way (movement of the working end to create the whole knot structure) but often the tying does not require pulling of the working end, or requires it only partially. I gave at the beginning an example with the shoelace. Textbooks often present several procedural variants for the most common knots such as the Bowline and the Clove Hitch [2].

It is important to observe that these variants are not easily predictable, and realization of the equivalence in their result often comes as a surprise (or, if not, as an interesting theorem).

(12) Metric knowledge. Knotters have an understanding of how much rope is needed to tie a knot (“Will it suffice?”, “You took too much/too little”)

(13) Handedness: Michel and Harkins [17] found that “observational learning of manual skills [knot tying] is significantly enhanced when the student and teacher are concordant in handedness”. Some video tutorials for knots present a subjective viewpoint on the hands, and those that do not may warn about the “mirror” effect created by looking at a video.

(14) People make systematic mistakes or encounter systematic difficulties in tying certain types of knot (eg. turning the final loop in tying a Cleat Hitch.)

(15) Generalizability. To some extent, once one has learned to tie a given knot, one can generalize (to thicker ropes, to specular knots, to different supports, to constrained tying, e.g. with a single hand). There are limits, though (once more, expertise is often context-bound). I learned a certain sequence for the Clove Hitch (“superpose rings in a “non intuitive way”), but this only holds for a rope's standing end that is presented on the right hand-side of the right hand. It is difficult for me to do the same for the left hand. I learned how to tie a Bowline with two hands, but I may need to do it with only one hand. I am better at tying the Cleat Hitch on the starboard side than on the port side. Under constraints (rain), I happened to have to tie a sheet to the roof using a Clove Hitch: no visibility, wet glasses, only one hand available, use of teeth, “generalizing”, starting from the memorized sequence, no visual control. It helped that I had memorized the sequence eyes shut. Success in some of these performances would speak in favor of some generality in mental representation of knots. More often than not, success does not appear to be at hand, thus indicating rigidity of the representation.

(16) Retrievalability: We have some understanding

of the knotting sequence, given perception of the final result of the knot. At some point I realized that a Clove Hitch is the result of tying two Half Hitches (Fig. 12).

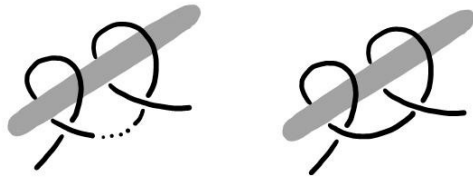


Fig. 12. Tying two Half Hitches in a sequence produces a Clove Hitch. Functionally, it is not a simple addition: the second step “closes” the Clove Hitch.

In the case at issue, the visual asymmetry of the final product masks the iteration within the sequence. We see the working end and the standing end “leave” in two different directions. But if we follow the movement of one of the two ends, we can appreciate the iteration.

Topologists can appreciate the compositionality of knots; a standard task is the decomposition of a knot into prime knots, i.e. knots that cannot be further decomposed. Once more, it is not obvious that this performance has an ecological counterpart. A topological knot can be decomposed into two trefoil knots; but a sequence of two plain knots on the same piece of rope does not automatically count as *one* knot.

Some imperfect understanding of compositionality may make one imagine impossible operations. I dreamed for a while of an “inverse” knot of a given knot, such that by combining the two and pulling ends I would end up with the unknot.

(17.1) Graphical competences 1. Drawing knots, given knowledge of a knot and of the knotting process (as opposed to copying a knot from life), is not trivial. Personal experience (Fig. 13) suggests that the best way to draw a knot (without copying it from life) is to retrace the movement that tied it. (Draftsmen who prepare drawings for manuals are likely to copy tied knots.)

(17.2) Graphical competences 2: deciphering diagrams. Over and above topological diagrams, diagrams are widely used in knot textbooks. Some diagrams appear to be more useful or more readable

and effective than others. Although it is difficult to provide a measure, we can point out some elements.

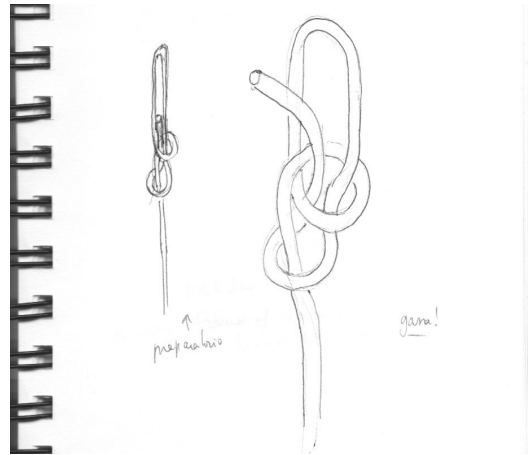


Fig. 13. Author's drawing of a Bowline from memory (left, preparatory sketch. 15.06.2012)



Fig. 14. The “little train” graphic method, based on the path followed by the working end.



Fig. 15. The “grab the bight” method. Circles, or hooks, indicate what to hold and where to move it in the next step.

The “little train” rendition method (Fig. 14) follows the topological construction of the knot and is in general of little assistance. The “grab the bight” rendition method (Fig. 15) models actual motor shortcuts that create the knot structure. Thumb rules for diagrams are derived by general indications about how to avoid cluttering graphic rendition, applied to the specifics of knots (Tufte [21]). Diagrams must represent intersections in order to convey the structure of the knot, thus intersection noise should be avoided. Tangents suggest intersections and are thus forbidden; information about intersections should be kept; irrelevant intersections ought to be avoided (Fig. 16). Graphic conventions about intersections (rope is not “cut” at the intersection, but is seen to continue under it) rely on Gestalt factors, such as the law of good continuation (“What is the continuation of what?”) which also underlies perception of physical knots.

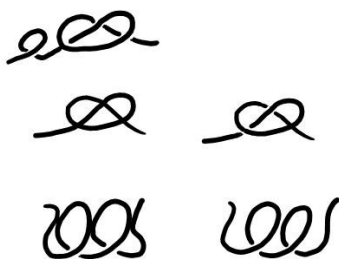


Fig. 16. On the left hand side column are examples of poor graphical renditions for knots. Top, an irrelevant loop is drawn. Middle, intersections are ambiguous. Bottom, tangents clog the image. Improvements appear in the right hand side column.

(18) We do not only categorize knots and have names for them; we also have names for knot parts. These are not only technical names such as 'bight',

'loop', 'elbow'. (See Fig. 1.). People use terms such as 'bunny ears' in teaching how to tie shoelaces (for a bight), or refer to the 'hole' or 'furrow' in describing a Bowline (for a loop). These terms, invented in order to name the parts of the knot or of the rope that contribute to knot structure, are metaphorical or analogical. The action repertoire for knotting, that includes complex interactions with rope and object, is fine-grained, and the scarcity of dedicated terminology is compensated for by metaphorical introductions.<sup>1</sup>

As we have seen, there is a large set of different but partly interdependent recognitional, practical and linguistic performances to be explained. What are the ingredients of the explanans?

### Knotting competence

Tying knots is a sequential action that uses repeated moves to create configurations of rope. Our first goal is to spell out the mental lexicon for the basic operations one performs on ropes when tying. Besides, knots are a wonderful case study for embodied and object-dependent cognition, as their realization depends on continuous object and sensorimotor feedback. The proper representation may involve not only the structure of the knot on a piece of rope, but the structure of the complex that includes rope and object.

In what follows I rely on Pignocchi [18] on the organization of the learning sequence of draftsmen. In learning to draw, children – but also adults – move from simple scrawls, the results of motor experiments, to more and more complex skeletal representations, that they are then able to integrate in images with an articulate content ( Fig. 17).



Fig. 17. Basic scrawls that merge into more and more complex graphic schemes in learning

<sup>1</sup> Semi-technical knot terminology (“working end/standing end”) appears to be recent; it is used by people who teach knots with words. As a contrast, the terminology of types of ropes used in sailing (mooring line, sheet, etc.) is probably older, as it is used to distinguish ropes with different functions.

Mastery of Atomic Graphic Schemes (AGS) controls the production of simple scrawls. Scrawls are represented as mentally undivided, accomplished in a single gesture. Once Atomic Graphic Schemes are stabilized, they can either be reused in more complex, Molecular Graphic Schemes (MGS), that are chains of AGS, or be slightly modified to fulfill other representational purposes. The repeated execution of MGS has in some cases the effect of making them to some extent automatic, thereby turning them into new, richer AGS.

The account has a number of theoretical advantages. For instance, it explains the difficulty in generalizing and the topic-boundedness of expertise. Draftsmen specialize: those who are good at drawing flowers are not thereby good at drawing faces. The theory also explains the peculiar stylistic traits of draftsmen, which depend on the idiosyncracies of AGS. The theory keeps “Darwinian” and “Lamarckian” aspects of creativity, introduced by Johnson-Laird, in balance. Little random variations in executing AGSs or MGSs may appear satisfactory and get stabilized by repetition (Darwinian aspect). General constraints on how to hold and move a pencil and on what counts as a representation control the exploration of new AGS (Lamarckian aspect). The account further predicts that at least some gesture that produced the drawing are perceptually retrievable.

The working hypothesis of the present article is that the theory of AGS provides a plausible model of knotting competence and of its development. Accordingly, one would learn some Atomic Knotting Schemes (AKS), reuse them in (compositional) Molecular Knotting Schemes (MKS) that, with practice, become or are treated like new atomic lexical entries. In learning an AKS, one associates a sequence of movements and a visual (or visuo-tactile) result. The peculiarities of learning, innovation, generalization and transmission would be explained by using the resources of the AKS-MKS framework. For instance, random variation in AKS can get stabilized by repetition; general constraints on how to tie knots control the search of new knots and condition the consolidation of MKS (balance between Darwinian and Lamarckian aspects). Finally, gestures behind knot production would be retrievable.

In developing AGS one relies on existing abilities. Holding an object like a pencil, or tracing a line in the sand with one's fingers, are proto-graphic

activities. Likewise, in the creation of AKS one relies on pre-existing abilities. Whoever has used a piece of rope (say, to walk a dog) knows that coiling it twice around one's hand renders the grip firmer. When we pull something, we often take advantage of fixed poles to reduce our effort. When coiling rope around a bar (e.g. around a tree) we easily discover that a Half Hitch configuration is extremely effective. These are proto-knotting activities and knowledge that get integrated in the simplest AKS. Creating MKS, on the other hand, involves the deployment of compositional abilities.

A side hypothesis is that one will, or will not, be able to tie an unknown knot by looking at the result, according to one's repertoire of AKS and MKS.

But what are the ingredients of AKS and MKS? In the following sections we describe some of the hypothetical ingredients of the mental computations involving AKS and MKS. Some of these are sub-atomic, such as the ability to generate and see certain relations between movements of the hand and configurations of rope. Others are of higher level, such as the chaining of AKS in long sequences.

### Parts of knots, of rope

Some parts of the rope become salient and are used as beacons for orienting the knotting process. The corresponding concepts may be lexicalized, or may be activated by analogies. In the traditional way to teach the Bowline, a segment is dubbed the “tree”, the working end is renamed “the rabbit”, and the loop is a “furrow” or a “rabbit's hole”.

*“A mnemonic used to teach the tying of the bowline is to imagine the end of the rope as a rabbit, and where the knot will begin on the standing part, a tree trunk. First a loop is made near the end of the rope, which will act as the rabbit's hole. Then the “rabbit” comes up the hole, goes round the tree right to left, then back down the hole.” (Bowline, Wikipedia entry, retrieved on 09.01.2012)*

Terminology in knotting practices is semi-technical (Fig. 18). It is not to be assumed that knotters know it, nor that it lexicalizes some mental concepts knotters have. Most likely it has been fixed by writers of knot books for teaching purposes.



Fig. 18. From left to right: working end, bight, loop, standing end.

### Loops and torsions

Asher [1] brings to the fore some important physico-geometrical properties of twisted rope. If you look along the axis of a piece of rope while twisting it e.g. clockwise, you will see that the rope undergoes a torsion; call this a right-handed torsion. Right-handedness is here an intrinsic, viewpoint independent characteristic of the rope (if you look at the rope from the opposite direction, it will appear to you right-handed as well.) In order to release the torsion, you can do either of two things: twist the rope counterclockwise, or coil it counterclockwise. If you look at the coiled rope along the axis of the coil, you will see that moving away from you the coil is left-handed.

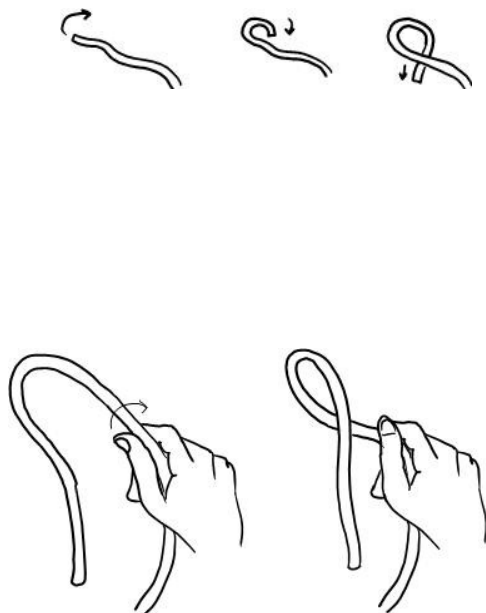


Fig. 19. Two ways to create loops.

Thus, you can produce a loop

- 1) by shifting the working end (Fig. 19, top).
- 2) By passing the thumb over the index finger to induce a torsion of the rope (fig 19, bottom).

A rope can be modeled as a series of rigid coaxial discs with a limited freedom of movement around their axis. The internal circumference of a loop is shorter than its external circumference. Each disk is then asked to rotate a bit in order to find room for the matter compressed in the internal part of the loop. Conversely, torsions automatically generate loops (Fig. 20).

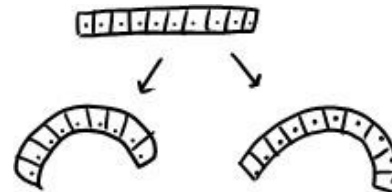


Fig.20 Ropes are things such that torsions generate loops and conversely.

These are physico-geometric properties of rope, that can be machine produced and machine measured. Two invariants surface: the right-handed torsion, that generates a left-handed coil, and the left-handed-torsion, that generates a right-handed coil (and conversely). As a consequence, the global shape of a part of the knot (of the coil) *stores some implicit information about the potential torsion of the rope.*

When learning to tie knots, a person performs twisting and coiling; these invariants are associated with sensorimotor primitives. Twisting rope provides haptic feedback. One *feels* the torsion, i.e. one feels that the rope tries to get back to its original shape. The tension at the tip of your fingers is *released* when a coil is formed. To form the coil, you just have to move your hands close to each other. Conversely, at the end of one coil, one realizes that one has generated a torsion, which can be eliminated by untwisting the rope. (Neglect of this operation has produced many a kinked water hose.)

To sum up, the basic rules are:

*Twisting and joining causes coiling*

*Coiling and separating produces twists.*

The basic knowledge of any knoter concerns the interaction of the physical and geometrical properties of rope. But although a knoter may implicitly know (feel) the torsion-to-loop interaction, she may be blind to the converse interaction. Beginners must be told that when coiling rope, for each loop they have to produce a torsion, otherwise loops will mess up.

Knowledge of knots is first and foremost storage in memory of these elementary sensorimotor regularities. The final *visual* shape of the coil is associated with a certain movement that produces or releases a torsion.<sup>1</sup>

A number of ecological knots are created by generating coils and making them interact. The Clove Hitch, when constructed in the middle of rope, without using the working end, is the result of the superposition of two coils. Creating coils is a prerequisite for executing these knots in an efficient way. We have seen that some most common knots are unknots, topologically speaking: the Clove Hitch, the Half Hitch. Let me add one more, less common knot to the lot, the Sheepshank (Fig. 21). The Sheepshank is the result of the pairwise intersection of three coils.

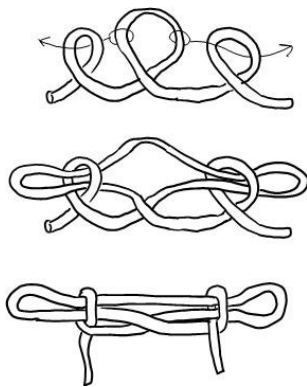


Fig. 21. The Sheepshank is actually an unknot, that only survives because of tensions.

These (un)knots, incidentally, have the advantage that they can be tied in the middle of rope, without access to the ends of the rope, by simple interaction

<sup>1</sup> The different ways of tying the shoelace knot described at the beginning generate different torsions, thus store energy in different ways in each knot. Accordingly, the generated knots are only superficially equivalent.

of loops. On top of loops/coils, one produces bights when executing a knot (by holding running and standing ends, each in one hand, and having the hands get closer to each other.)

### Pass-through

If some (un)knots require no access to the working end, “real” knots are in general tied by having the working end pass through a loop. A basic principle governs knot production.

*Working-end-and-loop axiom.* You can only tie a (real, i.e. non-unknot) knot by having the the working end pass through a loop.

This is a necessary but in no way sufficient condition, as you can tie unknots that way (e.g. the Clove Hitch, that can be tied directly on the standing end.)

### The role of visual crossing

Intuitively, no matter how many times you coil a piece of rope around a pole, you won't thereby have a knot. But coil once and cross over, and you'll have a Half Hitch. Crossing is related to twisting and coiling. The structure of knots involves a passage through a loop, whereby the principle that:

*the 2d projection of any knot will always involve a crossing.*

Thus perceptual crossing (the presence of an x-junction in the image) is a necessary condition for being recognized as a knot (it is not a sufficient condition, as the unknot can present crossings)

Crossing is in general an important condition in assessing topological equivalencies. The “tied” double donought is topologically equivalent to the “untied” double donought, notwithstanding their visual difference, that suggests a topological distinction. The perceived difference is an x-junction.

### Knowledge of the sidedness of a loop: guaranteeing stability

Another piece of intuitive knowledge concerns the interaction between the intrinsic orientation of the loop and the side of the loop from where the working



end must enter if one wants to get a stable knot structure (i.e. the knot will form and not collapse into the unknot.) When starting the Bowline or the Sheepshank or the Plain Knot, I know that the working end must enter from one, and not the other, side of the loop; or that a bight must pass on one, and not the other, side of the loop. I immediately see that the wrong side will not provide a stable structure. I know whether I am creating a “building step” or, sadly, an “empty step”. One has intuitions about what will work.

In tying a Bowline, one can compute the correct sidedness at each step. Knowing that the working end should exit the loop where it entered the loop, one must first determine the correct way to enter the loop, and consequently the correct way to create a bight around the standing end (Fig. 22).

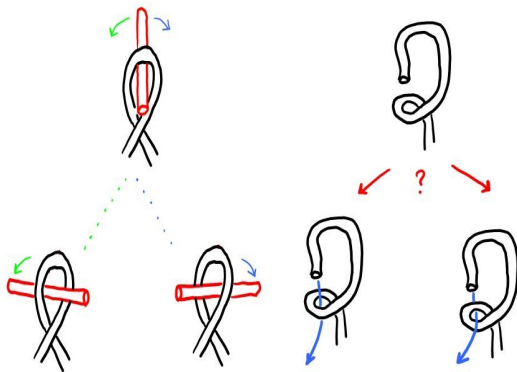


Fig. 22. Only the right bottom interaction of each case stabilizes the tension stored in the loop. (The stabilizer can be an external object, or a piece of the same rope used for the loop.)

### Knot and object: external dependency

Many knots are used to create ties between objects, or to fix rope on an object. This invites discussion of a complex set of invariances and, consequently, of sensorimotor contingencies. The first, basic principle concerns unknots:

Unknot dependency: *An unknot can be tied only on an object.*

Clove Hitches, Cleat Hitches, Half Hitches are unknots whose survival depends upon the existence of an object they are tied around. They are *dependent knots*.

There are seeming counterexamples: you can tie a Clove Hitch on a portion of the very same rope you are using for the hitch. Now, although the “local” movements are those for creating a Clove Hitch, the end result is a Plain Knot. The seeming counterexample allows us to distinguish two senses of dependency:

*self-dependency* (e.g. you can tie a Clove Hitch on the standing end) and

*other-dependency* (e.g. you tie the Clove Hitch on an object that is not the rope itself.)

### Two-object topologies

This introduces the theme of two-object topologies.<sup>1</sup> The formal counterpart of two-object topologies involving knots is the study of *links*. Once more, the descriptive gap between link theory and ecological links is as wide as the gap between knot theories and an account of ecological knots. The arguments are the same that we used for topological knots. For instance, topological links are not the result of tying, whereas ecological links are tied.

We should distinguish the metaphysical properties of self/other-dependency from the functional properties of self-reference. Some knots are self-referential: they are only used to store ropes, or to reduce the volume or length of rope.

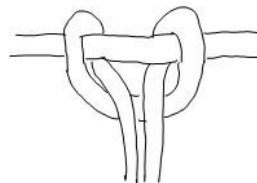


Fig. 23. A notoriously slippery knot, the Cow Hitch is used to “serve the rope”, fixing it so as not to create a mess, and allowing for quick untying.

Self-dependency and self-reference are orthogonal to each other. The Cow Hitch (Fig. 23) and the Sheepshank are self-referential in that they are used to change the properties of the rope and not for fixing anything. But the Cow Hitch is other-dependent, whilst the Sheepshank is self-dependent. (Functionally, the Sheepshank is used to shorten the

<sup>1</sup> Two-object topologies were used by Casati and Varzi [7] in order to account for some classificatory performances related to holed objects. Considering holes as completely filled, the topological properties of the contact surface between the host and the filler correlates fairly well with most commonsense categories of holes.

available portion of rope, or to use rope with a damaged portion.)

### Object-based functional lexicon

We have seen that the lexicon for knots introduces terms to characterize parts of the rope and elements of knots. The part of rope, or the external object that stabilizes the loop (see again Fig. 22) is a stabilizer. Whether a part of rope or an external object can act as a stabilizer depends on its capability to counter the tension created by the loop. In the case of rope, this may in turn depend on the stabilizer part's tension, or on its weight (for instance, if the standing end is long enough).

In view of the importance of external objects in tying knots, we need some semi-technical terminology for supports (Fig. 24).

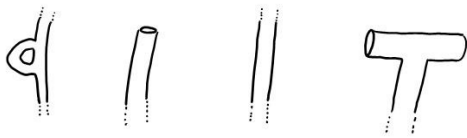


Fig. 24. Handlebody; pole; bar; cleat. Dashed lines indicate that the object's bounds are not within reach; the object is to be considered endless for knotting purposes (rope will not be allowed to slip out from there.) The cleat is topologically equivalent to the bar, but its shape conditions the knot. Starting from this simple taxonomy, endless compositional variations are available.

As we noticed, geometric features of knots are related to their causal properties (stop knots, stable knots, etc.) There is further an interaction of geometrical features of knots and physical properties of both ropes and things tied. For instance, one can create a figure-of-eight stop knot that is large enough so as not run through the handlebody. The elementary morphologies of Fig. 23 provide a lexicon of basic shapes. There is no upper bound to the complexity of object shapes one can use to create links. The basic lexicon helps characterizing the elementary interactions between rope and object. Tying a knot through a handlebody takes advantage of the topology of the handlebody to constrain the movement of the rope, and at the same time the handlebody requires that the working end passes through the hole in it. The pole is less constraining and at the same time allows for knotting procedures that do not involve the working end. Another relevant aspect of two-object interaction, in the case of knots, is that one of the two objects is in the norm

underformable. This means that one can take advantage of its rigidity in the execution of the knotting sequence. A final object feature knotting takes advantage of is the *permanence* of the topology of the object. We do rely on the fact that objects (as opposed to rope) do not change their topology. (And indeed, we are surprised when this happens, for instance when topological properties of the object can migrate to the rope. If I tie up my arms by crossing them, and then grasp two ends of a piece of rope with my hands, and then open up my arms, I end up with a Plain Knot on the rope, and no knot on my arms. The knot has *moved* from my arms to the rope.)

The features in question defy classification; geometry is intertwined with function. Topologically there is no difference between the pole and the cleat, and although there is no topological difference between Cleve and Cleat Hitch, as the execution is controlled by the object, it results in two utterly different procedures. Or, consider ring and bar (an example of a bar would be a tall tree, around whose trunk one ties a knot). Functionally they could be considered equivalent: their end segments do not exist or are not accessible, so that the tying procedure requires a use of the working end. But the ring has the further property of keeping the knot in a certain place.

### Knowledge of knots, the original loop, and the fundamental role of the Half Hitch

Half Hitches are ubiquitous components of knots (cf. once more Fig. 12, showing that the Cleve Hitch is the result of tying two Half Hitches in sequence. Now, the Half Hitch is an unknot: it is a simple loop. According to Unknot Dependency, it can be tied only on an object. The object acts as a stabilizer of the loop. This object – according to Self-Dependency, can be be another part of the same rope. This has in general the consequence of turning an unknot into a knot. The Half Hitch then “becomes” a Plain Knot. This elementary dynamics is at the basis of most knotting. The Half Hitch stabilizes the tension generated in the creation of the loop. Knowledge of knot is thus mastering of operations that orchestrate the management of the energy stored with the creation of the original loop.

### (Provisional) conclusions

We have seen that some ecological knots are, mathematically speaking, unknots, and thus that the

topological theory of knots is at best a partial account of knotting abilities. We have further seen that the mental representation of knots should allow for limited generalization, understanding of knot equivalences and compositionality. Even if you have some understanding of the compositional structure of tying a simple knot, you do not thereby have an access to the end result of just any knotting procedure that involves the elements of the composition. The consequences of the atomic actions you perform are not easy to predict; not even for experts. As knotting involves external objects essentially, the feedback loop that unites perception and action is essential for our understanding of them; knotting provides an ideal case for situated cognition and externalized mental procedures. It is early to provide a formal characterization of the principles at play – an algebra of knotting and knot understanding, as if it were, as opposed to an algebra of knots. Some existing models accounting for motor-perceptual performances (e.g. models for drawing) can be reused in the case of knotting, thus allowing us to distinguish an atomic level, with sub-atomic parts, and a molecular level. We were able to enlist some principles at the atomic or subatomic level, all involving sensorimotor representations: the relevance

of parts of rope and their lexicalization; the duality of twisting and coiling and the contribution of both to the storing of action into the configuration of the knot; the interaction of loops and ends to create stable structure (good and bad 'sides' of the loop) and the consequence for the visual aspect of the knot, that must include x-junctions; the interactions of rope with object shape and topology and the lexical saliency of functional object features; the necessity of dealing with two-object representations; the object-dependency of unknots; the distinction between self-dependency and self-reference; and the fundamental role of loop stabilization in half hitches, that turns out to be the most important subatomic elements of knots.

The present article pleads for the investigation of process topology as opposed to static topology. Shapes are usually considered as static properties. But in the case of ecological knots, their features bear a trace of the process that led to them, that included planning, motor execution, and perceptual control, in the service of the management of the energy stored in the shape of a rope to create stable structures.

## References

- [1] Asher, H., 1989, *The alternative knot book*. London, Nautical Books.
- [2] Ashley, C.W., 1944, *The Ashley Book of Knots*. New York: Doubleday.
- [3] Balari, S., Benítez-Burraco, A., Camps, M., Longa, V.M., Lorenzo, G., Uriagereka, J., 2011. The archaeological record speaks: Bridging anthropology and linguistics. *International Journal of Evolutionary Biology* 2011, doi:10.4061/2011/382679.
- [4] Balari, S., Benítez-Burraco, A., Camps, M., Longa, V.M., Lorenzo, G., 2012, Knots, Language, and Computation: A Bizarre Love Triangle? Replies to Objections. *Biolinguistics*, Vol 6, No 1, 79-111.
- [5] Bayman, 1977, Theory of hitches, *American Journal of Physics*, 45 (1977) 185
- [6] Camps, M., Uriagereka, J. 2006. The Gordian knot of linguistic fossils. In J. Rosselló & J. Martín (eds.), *The Biolinguistic Turn. Issues on Language and Biology*, 34–65. Barcelona: Publications of the University of Barcelona.
- [7] Casati, R., Varzi, A.C., 1994, *Holes and Other Superficialities*. Cambridge, Mass. MIT Press.
- [8] Casati, R. 2009. Does topological perception rest on a misconception about topology? *Philosophical Psychology* 22 (1):77 – 81.
- [9] Chen, L., 2005 The topological approach to perceptual organization. *Visual Cognition*, 12, 553-637.
- [10] Cross, E.S., Cohen, N.R., Hamilton, A.F.d.C., Ramsey, R., Wolford, G., Grafton, S.T., Physical experience leads to enhanced object perception in parietal cortex: Insights from knot tying. *Neuropsychologia*, 50, 3207-3217.
- [11] Crowell, B. undated, *Knot Physics*. <http://www.lightandmatter.com/article/knots.html> , last retrieved Feb 01, 2013.
- [12] De Toffoli, S., Giardino, V., 2013, Forms and Roles of Diagrams in Knot Theory. Ms.
- [13] Herzfeld, C., Lestel, D., **Knot tying in great apes: etho-ethnology of an unusual tool behavior** *Social Science Information* 2005 44: 621-653.
- [14] Karmo, T., 1977, 'Disturbances', *Analysis* 37: 147-148.
- [15] Lobina, D.J., 2012, All tied in knots, *Biolinguistics*, 6, 1, 70-78.
- [16] Lobina, D.J., Brenchley M., 2012, Language, Knots and Computation: More Bermuda than Love. *Biolinguistics*, 6, 2, 176-204.
- [17] Michel, G. F., Harkins, D.A. 1985. Concordance of handedness between teacher and student facilitates learning manual skills. *Journal of Human Evolution* 14, 597–601.
- [18] Pignocchi, S., 2010, How the intentions of the draftsman shape perception of a drawing. *Consciousness and Cognition*, 19(4):887-98.
- [19] Sossinsky, A., 2002 *Knots: Mathematics with a twist*. Harvard University Press.
- [20] Strohecker, C., 1999 *Why Knot?* PhD thesis, MIT.
- [21] Tufte, E.R., 1997, *Visual Explanations: Images and Quantities, Evidence and Narrative*, Cheshire, CT, Graphics Press.

- [22] Tracy, J., Flanders, A., Madi, S., Laskas, J., Stoddard, E., Pyrros, A., Natale, P., DelVecchio, N., 2003, Regional Brain Activation Associated with Different Performance Patterns during Learning of a Complex Motor Skill., *Cerebral Cortex*, 13:904-910.
- [23] Urton, G., 1998 "From Knots to Narratives: Reconstructing the Art of Historical Record Keeping in the Andes from Spanish Transcriptions of Inka *Khipus*". *Ethnohistory* 45 (5): 409–438.

### **Acknowledgments**

I'm indebted to the members of the perception seminar at Institut Nicod, and to Achille Varzi, Alessandro Pignocchi, Alessandro Scorsolini, Paolo Biagini, Andrea Formica, Paulo Santos, Valeria Giardino and Silvia De Toffoli for discussion. Errors are my sole responsibility.



# Graph-based pattern recognition and applications

Roberto Marcondes CESAR JR.<sup>1</sup>  
*Universidade de Sao Paulo, Brazil*

**Abstract.** Structural pattern recognition plays a central role in many applications. Recent advances include new theoretical results, methods and successful applications. In the present talk, some recent graph-based methods for shape analysis will be shown. The presented methods include a new representation for graph-matching-based interactive segmentation and models for the analysis of spatial relations between objects. Applications will be presented and discussed.

---

<sup>1</sup>E-mail: cesar@ime.usp.br



# On Shape Poems and the Shape of Poems

## *A Computational Creativity Perspective*

Simon COLTON

*Computational Creativity Group  
Department of Computing, Imperial College, London  
ccg.doc.ic.ac.uk; sgc@doc.ic.ac.uk*

The Painting Fool is a computer program that we hope will one day be taken seriously as a creative artist in its own right. It has been developed within a Computational Creativity paradigm to exhibit behaviours which could be called skilful, appreciative and imaginative. At its heart is a two stage process of constructing a representative scene consisting of a set of coloured shapes and then simulating artistic media and techniques in rendering the shapes into a painting.

For scene generation, we have employed various graphics, machine vision and artificial intelligence techniques in order to simulate imaginative visualisation. These techniques include image segmentation, image filtering, emotion detection, constraint solving, context free grammars and machine learning. The methods are combined using a layering approach whereby shapes from one layer can be substituted by others in various ways.

When rendering an image, the software simulates the usage of a tool such as a pencil, pastel, crayon or paintbrush applying pigments to a medium such as a sheet of paper or a canvas. It has the ability to trace the outline of shapes and fill shapes in various ways. It paints in layers and can change the style in the layers at the shape level and the painting level to simulate the production of multimedia pieces.

We have written papers and given exhibitions which raise some philosophical issues related to the building of autonomously creative software. These include how to simulate imagination, emotional response and intentionality in software, the value of programs framing their output, how to evaluate the creativity of software, and what it means to hand over creative responsibility to a machine.



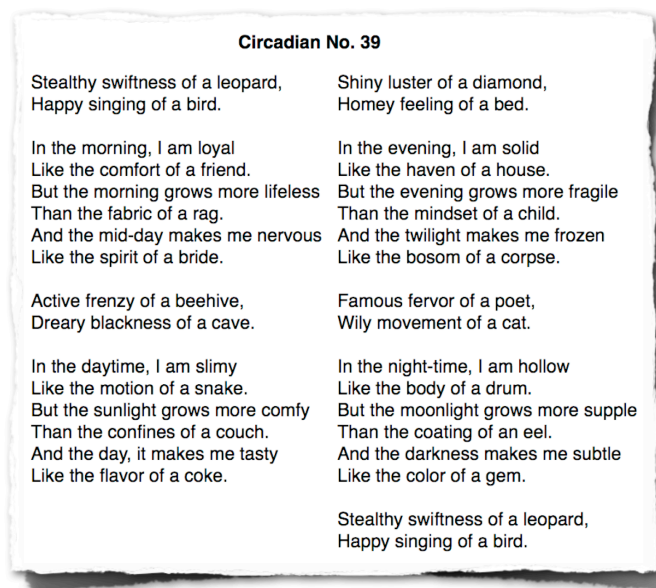
We have recently given The Painting Fool abilities to generate text artefacts such as poems and commentaries. Two example poems from a recent paper are given in figure 2, with the latter given with an automatically generated commentary about how it was produced. The first is clearly in a poetic form, while the latter has less strict shape as a poem.

In other work, we have enabled the software to take any font and plot paths through the centre of each letter shape, so that they can be drawn using the simulated pencils, pastels and paints available. This has allowed the usage of text in the paintings that The Painting Fool produces, and for the (future) production of multimedia pieces such as illustrated poems and collages.

In our most recent work we have enabled word wrapping so that text can be fitted into an arbitrary shape produced by the software or dictated by a user as per the shapes in this paper. This is done in a two step process whereby the size of the font is first searched for then the best arrangement of the words at that size is searched for.

The new abilities have led to the production of new pieces of art, as portrayed in figures 2 to 5. These include visual artworks where a shape is picked out by text and where a series of text shapes is used to depict something representational such as a portrait. The Painting Fool can now also produce shape poems, where the shape of the text somehow represents the subject material.

These new projects which combine text and shape for visual art and poetic purposes have highlighted a number of higher level issues in Computational Creativity. These include programming imperfection (the productive use of bugs), the latent heat issue (handing over creative responsibility often leads to lower quality artefacts being produced) and the benefit of getting software to frame its work with commentaries, titles and other texts.



**Figure 1.** Two poems produced by The Painting Fool.

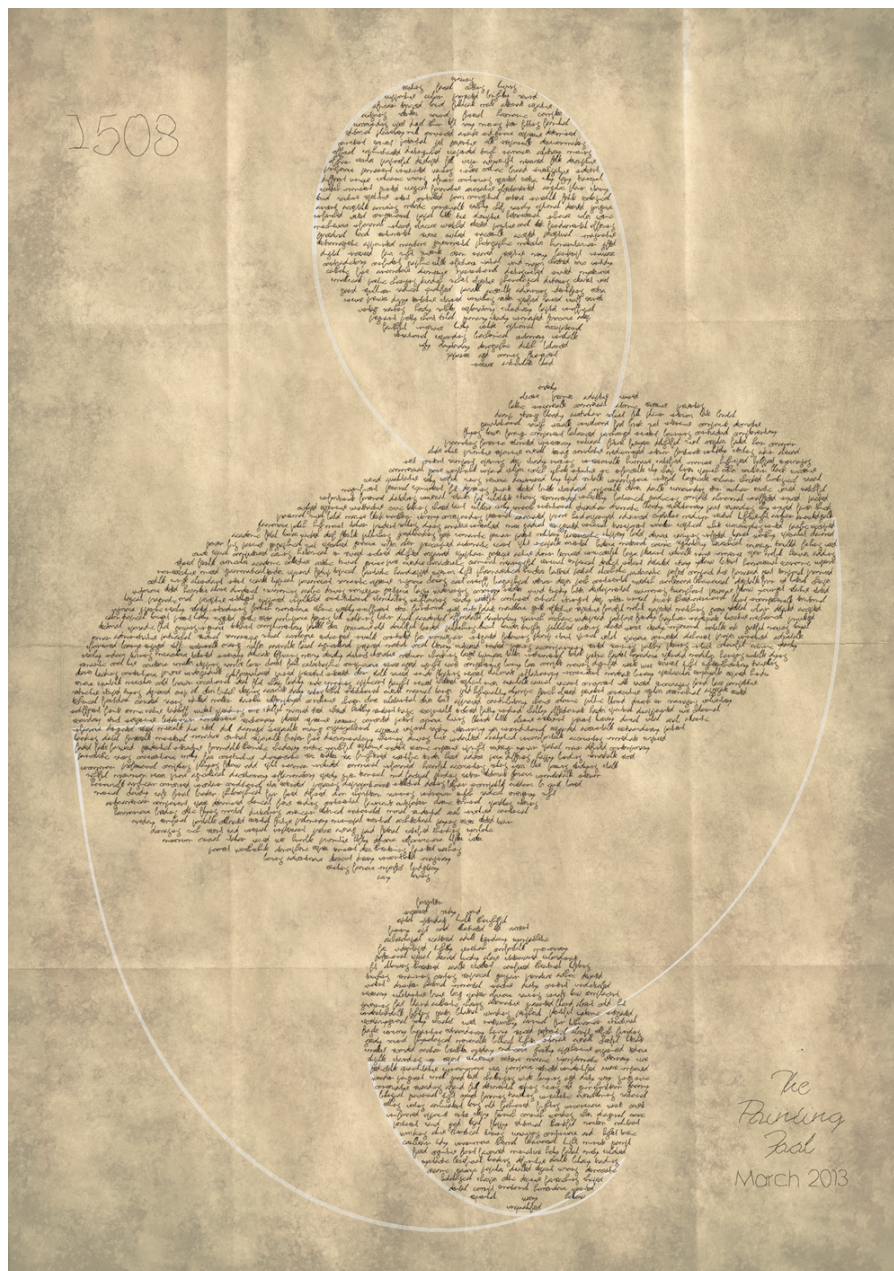


Figure 2. Artwork produced by The Painting Fool where an individual shape is highlighted by text.



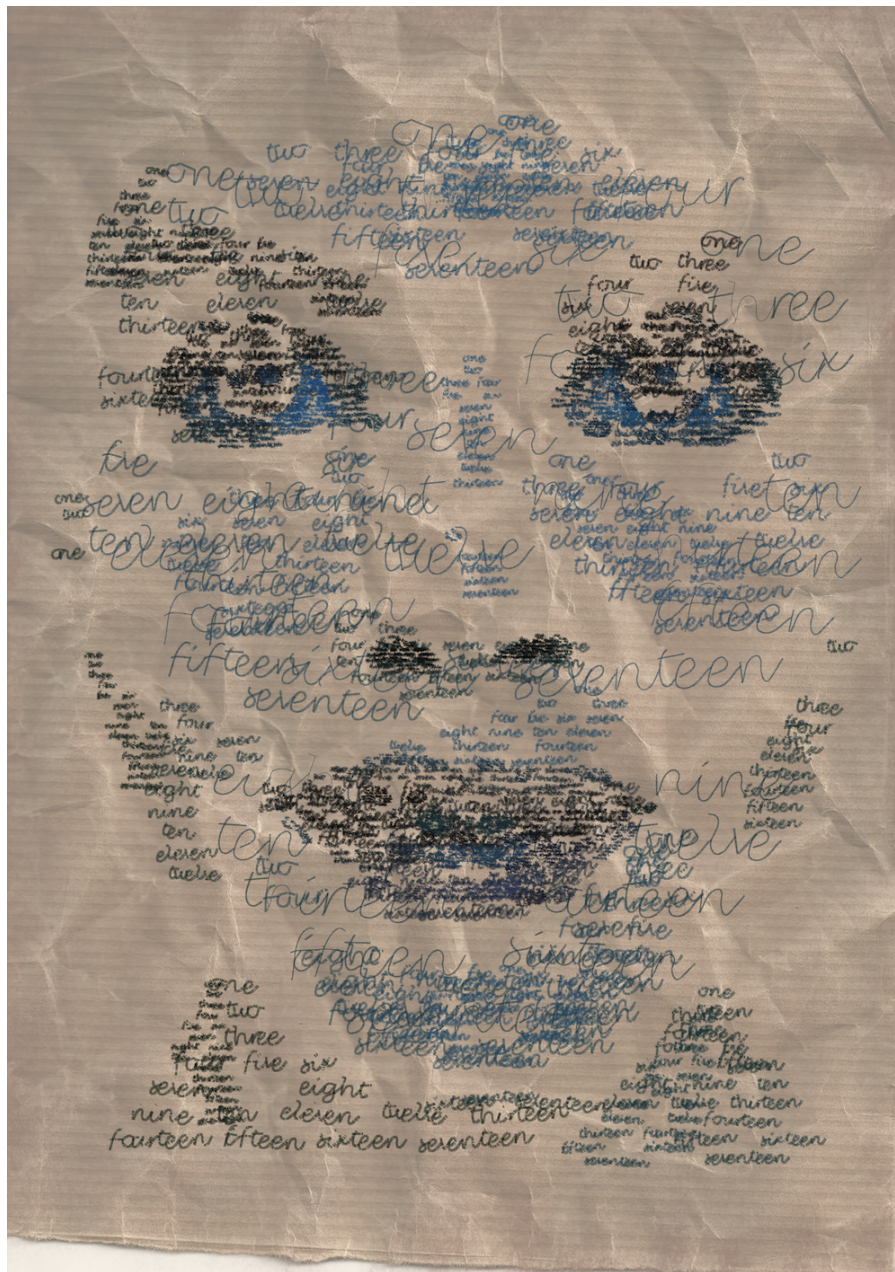


Figure 3. Artwork produced by The Painting Fool with a set of text shapes producing a portrait.



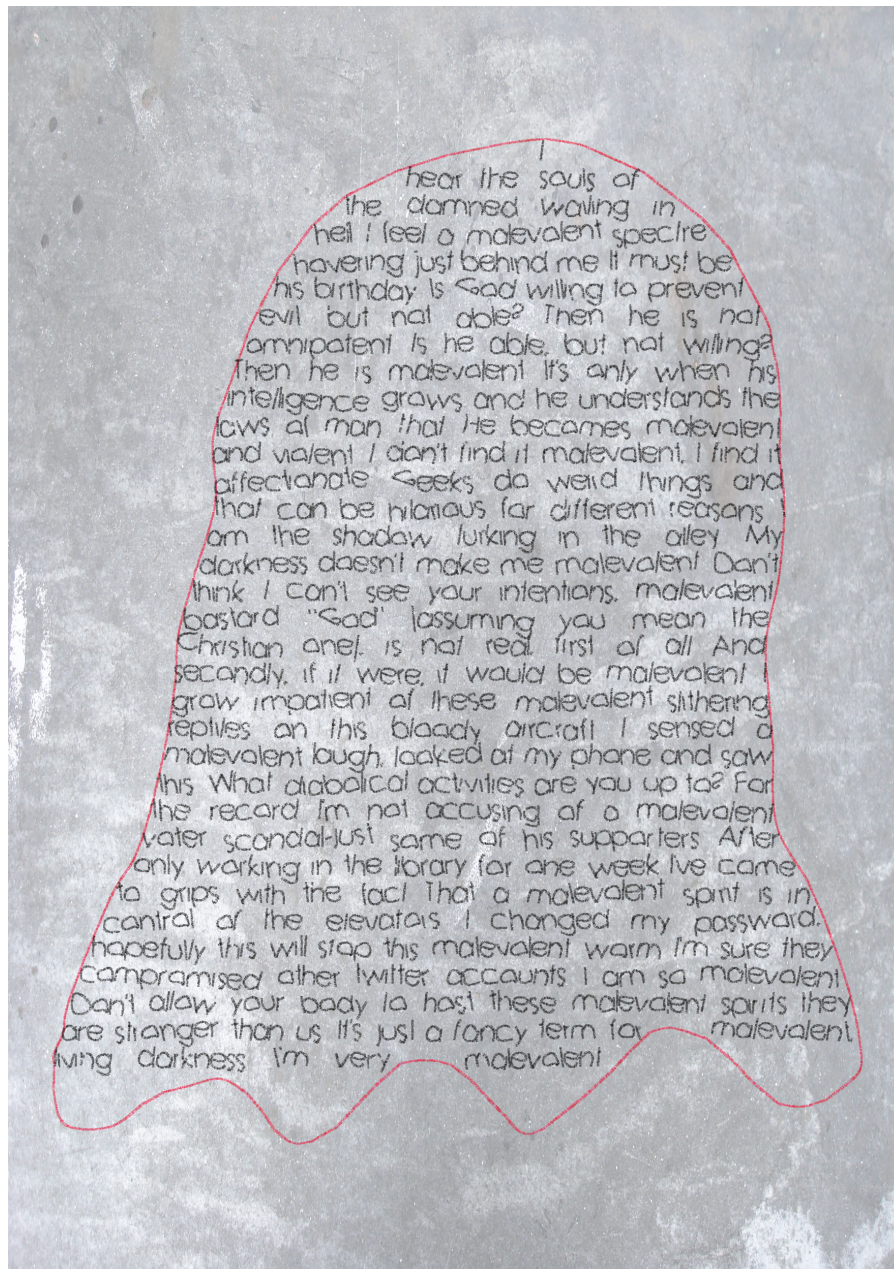


Figure 4. The Malevolent Ghost shape poem by The Painting Fool. Sentences from Twitter.

# Prolegomena to an Ontology of Shape

Antony GALTON

*College of Engineering, Mathematics and Physical Sciences  
University of Exeter, UK.  
Email: apgalton@ex.ac.uk*

**Abstract.** Influenced by the four-category ontology of Aristotle, many modern ontologies treat shapes as accidental particulars which (a) are specifically dependent on the substantial particulars which act as their bearers, and (b) instantiate accidental universals which are exemplified by those bearers. It is also common to distinguish between, on the one hand, these physical shapes which form part of the empirical world and, on the other, ideal geometrical shapes which belong to the abstract realm of mathematics. Shapes of the former kind are often said to approximate, but never to exactly instantiate, shapes of the latter kind. Following a suggestion of Frege, ideal mathematical shapes can be given precise definitions as equivalence classes under the relation of geometrical similarity. One might, analogously, attempt to define physical shape universals as equivalence classes under a relation of physical similarity, but this fails because physical similarity is not an equivalence relation. In this talk I will examine the implications of this for the ontology of shape and in particular for the relationship between mathematical shapes and the shapes we attribute to physical objects.

**Keywords.** shape ontology; mathematical vs physical shape; intrinsic vs embedded shape

## 1. Introduction

What are shapes, and how are shapes related to things which are not shapes? Are there indeed such things as shapes at all, entities of some sort that have a place in an inventory of the world's contents? Or can we explain talk about shapes in terms of an ontology in which shapes do not feature as entities of any kind?

There seem to be two distinct kinds of shapes: *physical shapes*, which we encounter in the physical world as the shapes of entities that exist in space, and *mathematical shapes*, which we encounter in geometry, the shapes of abstract mathematically-defined constructions. In both cases it seems evident that shapes are *ontologically dependent* on the objects whose shapes they are (their *bearers*), but the relationship between the two kinds of shape is not necessarily straightforward. I shall defer till later a discussion of mathematical shapes, and for the moment concentrate on physical shapes, the shapes of physical objects.

Granted, then, that physical shapes are always shapes *of* things, what kinds of things *have* shapes in this sense? A brief catalogue might be as follows:

1. *Material objects*, including chunks of matter (e.g., a pebble), organisms (e.g., a penguin), and assemblies (e.g., a bicycle).
2. *Non-material physical objects* such as holes, faces, and edges.
3. *Aggregates* such as a flock of birds or a cluster of buildings.

The boundary between material objects and aggregates is not sharp, since even a chunk of matter is, at submicroscopic resolution, an aggregate of atoms. With aggregates it is often not easy to determine an exact boundary, or therefore an exact shape [5], and to the extent that objects may be similarly indeterminate, it may likewise be impossible to assign exact shapes to them. I shall take this issue up later in the discussion of shape approximation. We should not assume uncritically that all material objects have shapes; Stroll [13] suggests that not all material objects have surfaces, and possession of a surface seems to be strongly associated with possession of a shape, even if not a necessary condition for it.

Phrases of the form “the shape of  $X$ ” and “ $X$  has such-and-such a shape” attest to the intimate relation between an object and its shape, characterised as an ontological dependence of the latter on the former. Other key elements of the object–shape relationship, to be accounted for in an ontology of shape, include two objects having *the same shape*, and an object *changing shape*, expressed using the sentence forms:

- (1)  $x$  and  $y$  have the same shape at time  $t$
- (2)  $x$  changed shape between times  $t_1$  and  $t_2$

In what follows, we will pay careful attention to these notions.

## 2. The dependency of shape upon objects

Shape-words in language typically come both as nouns and adjectives: in English we have, alongside nouns such as “circle”, “triangle”, “sphere”, and “cylinder”, the respective adjectives “circular”, “triangular”, “spherical”, and “cylindrical”. Cases in which the nominal and the adjectival functions are borne by the same word—e.g., “square” and “oblong”—are the exception rather than the rule. We also freely form compound adjectives such as “pear-shaped” and “heart-shaped”, and in some cases the noun forming the first part of the compound refers, not to the physical object which it normally designates, but to some mathematical shape conventionally abstracted from it—e.g., the “heart” symbol  $\heartsuit$  only very approximately resembles the complex three-dimensional shape of an anatomical heart.

Shape adjectives point to the notion of shape as a *property* of objects, whereas shape nouns point to shapes as entities in their own right. Which of these two pictures enjoys logical or ontological primacy over the other, and how are the two pictures related?

Ontological parsimony suggests that shape-as-property should take priority over shape-as-entity. Looking around us, we see physical objects, each with its own shape, but to suggest that we see the shapes *as well as* the objects smacks of ontological over-abundance. It is more natural, when in a parsimonious mood, to say that each object is shaped in such-and-such a way, where this notion is

expressed using a shape adjective. Thus we can say that the table is square, rather than that it stands in some relation to a shape entity which is *a square*.

A logical analysis of this view will invoke *shape predicates*, leading to predications of the form  $Square(x)$  or  $Circular(y)$ —or rather, allowing for the fact that objects can change shape,  $Square(x, t)$ , etc. A major disadvantage lurks behind the attractive simplicity of this scheme: if we want to generalise over shapes, we have to quantify over predicates, and this requires the use of second-order logic, with all the difficulties that that brings in its wake. Thus to express the sentence-forms (1) and (2) we would have to write something like

$$\begin{aligned} (1a) \quad & \forall \Phi (ShapeProperty(\Phi) \rightarrow (\Phi(x, t) \leftrightarrow \Phi(y, t))), \\ (2a) \quad & \exists \Phi_1 \exists \Phi_2 (ShapeProperty(\Phi_1) \wedge ShapeProperty(\Phi_2) \wedge \\ & \Phi_1(x, t_1) \wedge \Phi_2(x, t_2) \wedge \neg \Phi_1(x, t_2) \wedge \neg \Phi_2(x, t_1)). \end{aligned}$$

A standard way of reducing such second-order predications to first-order form is by *reifying* the properties expressed by the predicates which are being quantified over [4]. By this means we introduce terms designating shape entities, and introduce a first-order *HasShape* predicate to relate objects to the shapes that they have: thus instead of  $Square(x)$ , say, we would write  $HasShape(x, square)$ . In effect, this is to accord priority to shape-nouns over shape-adjectives. Our sentences (1) and (2) can now be expressed in first-order form as

$$\begin{aligned} (1b) \quad & \forall s (HasShape(x, s, t) \leftrightarrow HasShape(y, s, t)), \\ (2b) \quad & \exists s_1 \exists s_2 (HasShape(x, s_1, t_1) \wedge HasShape(x, s_2, t_2) \wedge \\ & \neg HasShape(x, s_2, t_1) \wedge \neg HasShape(x, s_1, t_2)) \end{aligned}$$

On this view, it is natural to regard shapes—that is, the entities designated by the  $s$  variables in (1b) and (2b)—as *generically dependent* entities. They are dependent, since a shape only exists insofar as it has bearers, and this dependence is generic because a shape is not dependent on the existence of a unique bearer but can be multiply realised in different bearers having the same shape.

Modern information systems ontologies such as BFO [7] and DOLCE [9] do not take this line; instead, they treat an object’s shape as *specifically dependent* on that object, meaning that the shape belongs uniquely to that object and cannot be shared with any other. In DOLCE, shapes, along with such things as colours, volumes, weights, and densities, are classified as *qualities*. The identity of an object’s shape is tied to the identity of the object itself: the shape comes into existence when the object comes into existence, and endures for as long as the object does. This does not mean that an object cannot change shape, though; what happens, according to DOLCE, is not that the object assumes a different shape, but that the object’s shape assumes a different *value*. The values that may be assumed by a quality are entities of another kind, called *qualia*, which collectively constitute a domain known as a *quality space*—in the case of shape, we could speak of *shape qualia* in *shape space*. These quality spaces are similar to the *conceptual spaces* of Gärdenfors [6].

On this picture, variability of shape shows up as a time-dependency, not of the shape on its bearer, but of the value of the shape on the shape. Writing  $shape(o)$  to refer to the shape which uniquely inheres in the object  $o$ , we have the rule



$$\text{shape}(x) = \text{shape}(y) \rightarrow x = y,$$

and our formulae now come out as:

- (1c)  $\text{value}(\text{shape}(x), t) = \text{value}(\text{shape}(y), t)$
- (2c)  $\text{value}(\text{shape}(x), t_1) \neq \text{value}(\text{shape}(x), t_2)$

The shape-as-quality view embraced by DOLCE has a solid pedigree in the Aristotelian four-category ontology that is encapsulated in the ontological square [8,12], which presents a cross-classification of the entities of an ontology along the dimensions of universal *vs* particular (distinguishing types from their instances) and substance *vs* accident (distinguishing independent from dependent entities). Thus *the roundness of this ball* is an accidental particular inhering in (and thus dependent on) the substantial particular *this ball*; and these two particular entities are instances of the accidental universal *roundness* and the substantial universal *ball* respectively. The ball itself is said to *exemplify* roundness.

The roundness of this ball is not quite the same as the shape of the ball conceived as a quality in DOLCE. The former is a *trope*, i.e., a specific instance of a property inhering in an object. In DOLCE terms, the “property” in question is not just a quality but a quality’s having a particular value. Thus a trope could be regarded as a quality/value pair. If the quality changes value (e.g., an object changes shape), then the previous trope is superseded by a new one. When the value of some quality changes continuously, there is a continuous succession of different tropes.

### 3. The primacy of “same shape” over “shape”

The reified analyses discussed above are predicated on the assumption that there are such things as shapes, whether universals or particulars, with a bona fide existence that must be accounted for by according them a place within our ontology. This can, however, be questioned. Consider again the two main ways in which we describe the shape of an object:

1. Using a descriptive adjective such as “square”, “round”, or a combination of adjectives such as “long and thin”;
2. By means of a comparison with some other object whose shape is assumed known, e.g., “heart-shaped”, “hourglass-shaped”.

In neither of these cases is there an explicit reference to shapes per se: we can understand “square” as a descriptive adjective without having to postulate any entity that is a square shape distinct from the square object we are talking about; and in saying that something is heart-shaped we are not saying that its shape is a heart, but rather that it is similarly-shaped to a heart. We do, of course, use the expression “it has the same shape as a heart”, which seems to suppose the existence of the shape as something distinct from the heart, but it may be argued that this too is a misunderstanding, the locution “has the same shape as” being more correctly paraphrased as “is shaped the same as”.

This way of arguing has venerable roots. It presents shape as one of a group of concepts *X* for which the notion of *X* itself is logically dependent on a prior

notion which, once we have the concept  $X$  at our disposal, it is natural to express using the words “has the same  $X$  as”. This latter notion is an equivalence relation which can be defined without any reference to the concept  $X$  itself. This idea is due to Frege [2], who noted that the concepts of *number*, *direction*, and *shape* can all be derived in this way.

In the case of (cardinal) number, the relevant relation is defined as follows:

- Set  $X$  has the same number as set  $Y$  if and only if there is a bijection (i.e., an exhaustive one-to-one correspondence) between the elements of  $X$  and the elements of  $Y$ .

Notice that bijections can be defined without reference to number; but on the other hand, according to Frege’s argument, number cannot be defined without having the prior notion of “same number” to establish an identity criterion for the new concept. Thus “same number” is shown to be logically prior to “number”. Instead of “has the same number as”, we can use the term “equipollent”.<sup>1</sup> We now define number in terms of equipollence as follows: The number of set  $X$  (i.e., the number of its elements, its *cardinality*) is the set of all sets equipollent to  $X$ .<sup>2</sup> This set is what we would now recognise as an equivalence class under the equipollence relation.<sup>3</sup>

Similarly, “direction” is logically dependent on the relation “has the same direction as”—which we routinely express as “is parallel to”—, and “shape” is logically dependent on “has the same shape as”, i.e., is geometrically similar to. In particular, the shape of an object can be defined as the equivalence class comprising all objects which have the same shape as it. This definition works well so long as (a) a domain of “objects” is established for the relation to be defined on, and (b) within this domain “same shape” can be defined as an equivalence relation. In the next section I consider as candidate domains, first, geometrical constructions, and second, physical objects; I then go on to consider what it means to say that a physical object has the same shape as a geometrical construction.

#### 4. Definition of the “same shape” relation

At the end of the previous section I glossed the relation “has the same shape as” as “is geometrically similar to”. The latter relation, however, is first and foremost defined as a relation on geometrical objects—which, for the moment, we may understand, in standard mathematical fashion, as subsets of  $\mathbb{R}^n$ , for some  $n \in \mathbb{Z}^+$ . We therefore need to ask in what way this relation can be applied to the very different domain of physical objects: very different because physical space is not a set of real-number triples, the usefulness of  $\mathbb{R}^3$  as a model for physical space

<sup>1</sup>Frege’s term was *gleichzählig* i.e., “equal-numbered”.

<sup>2</sup>Frege did not himself formulate this in terms of sets: he spoke of the number which belongs to the concept  $F$  (*die Anzahl, welche dem Begriffe F zukommt*) and equated this to the extension of the concept “equipollent to the concept  $F$ ” (*der Umfang des Begriffes “gleichzählig dem Begriffe F”*).

<sup>3</sup>This is not unproblematic: What set is this relation defined on? Frege supposed this could be *the set of all sets*, but as Russell pointed out to him, this notion leads inexorably to devastating paradoxes. An adequate discussion of this point would take us well out of scope of this paper.

being rather that it affords constructions which capture at least some parts of the abstract essence of phenomena in physical space that we wish to model.

#### 4.1. Similarity of geometrical objects

Considering first the notion of geometrical similarity as it applies to objects in geometrical space, the key notion is that of *distance*, which serves as a measure of the separation between two points. Writing  $\Delta(p, q)$  for the distance between points  $p, q \in \mathbb{R}^n$ , defined by the usual Pythagorean rule, we have:

**Definition of geometric similarity between figures in Euclidean space.**<sup>4</sup> Two subsets  $X$  and  $Y$  of  $\mathbb{R}^n$  are geometrically similar if and only if there is a bijection  $\phi$  from the points of  $X$  to the points of  $Y$  such that, for some constant  $\kappa \in \mathbb{R}^+$ , the following relation holds:

$$\forall x, x' \in X. \Delta(\phi(x), \phi(x')) = \kappa \Delta(x, x').$$

In other words, distances between points in  $X$  are multiplied by a constant factor  $\kappa$  when the points are mapped by  $\phi$  into their images in  $Y$ . This is straightforward and familiar. It is of particular importance to note that the relation thereby defined is an equivalence relation, and it is this that enables the Fregean move by which the shape of a figure can be identified with the equivalence class of all figures having the same shape as it.

#### 4.2. Similarity of physical objects

When we turn from  $\mathbb{R}^n$  to the physical world, things are less straightforward. Whereas distance in  $\mathbb{R}^n$  can be defined mathematically, in physical space the notion of distance is inextricably tied up with that of *measurement*, and the key fact about measurement here is that all measurement has finite precision. This means that whereas in  $\mathbb{R}^n$ , since distances can be arbitrary non-negative real numbers, the space of possible distances is simply  $\mathbb{R}^+ \cup \{0\}$ , the space of measured distances in physical space cannot take this form. To see this, note that we cannot meaningfully ask whether the length of a rod in metres is rational or irrational.

Given that in physical space we can only characterise distances in terms of measurement, and that measurement always has a finite precision, corresponding to the resolving power of the measuring instrument, it follows that geometrical similarity for physical objects can only be defined relative to a specified level of resolution. Consider two objects whose shapes we wish to compare, say  $P$  and  $Q$ , where  $Q$  is at least as big as  $P$ . Suppose the volume of  $P$  is  $v$  and that the resolving power of our measuring instrument is such that the smallest distance we can distinguish is  $h$  (I shall describe this as “resolution  $h$ ”). Then in principle, within the physical space occupied by  $P$  we can distinguish, say,  $n \approx v/h^3$  points,

---

<sup>4</sup>Here I am only dealing with Euclidean space—complications arise when we turn to non-Euclidean spaces. For example, on the surface of the sphere, figures cannot be similar without also being congruent. This is because, in this space, the sum of the interior angles of a triangle exceed  $2\pi$  by an amount that is proportional to the area of the triangle, and hence no figure can be expanded or contracted without changing shape. Thus only in Euclidean space is shape completely independent of size.

and to each pair  $x, y$  of these points we can assign a distance  $\Delta_h(x, y)$  that is some multiple  $\kappa h$  of the minimum discernible distance.<sup>5</sup> Let  $S_h(P)$  be this set of  $n$  discernible points in  $P$ ; we may think of them, if we wish, as “blobs” of diameter  $h$ , though this is not really how they seem to us as observers.

To say that objects  $P$  and  $Q$  have the same shape is to say that the points we can discern in  $P$  at resolution  $h$  can be mapped onto some set of points we can discern in  $Q$  at that resolution, such that, first, the distances between pairs of the latter set of points are not discernibly different, at resolution  $h$ , from some constant multiple of the distances between the pairs of points from the former set to which they correspond under the mapping; and second, *every* point in  $Q$  discernible at resolution  $h$  is “sufficiently near” one of the points corresponding to a discernible point in  $P$ . In other words:

**Definition of “same shape” for physical objects.** Physical objects  $P$  and  $Q$  (where  $Q$  is at least as big as  $P$ ) have the same shape, at resolution  $h$ , if, for some constant  $\kappa \geq 1$ , the set  $S_h(P)$  of points discernible in  $P$  at resolution  $h$  can be mapped into the set  $S_h(Q)$  of such points of  $Q$  by means of an injective mapping  $\phi$ , such that the following relations hold:

1.  $\forall x, y \in S_h(P). |\Delta_h(\phi(x), \phi(y)) - \kappa \Delta_h(x, y)| \leq \frac{1}{2}h$
2.  $\forall x \in S_h(Q). \exists y \in S_h(P). \Delta_h(x, \phi(y)) \leq \frac{1}{2}\kappa h$

This is perhaps as near as we can come to the notion of geometric similarity in a physical setting, in which the idea of “exact distance” gains no purchase.

An immediate consequence of this is that a pair of objects which come out as having the same shape at one level of resolution may have different shapes at a finer level of resolution. For physical objects, the concept of “same shape” is inescapably tied to the level of resolution at which the objects are examined; and since, according to the Fregean argument, the concept of “shape” is logically dependent on the concept of “same shape”, it follows that the notion of shape is also tied to levels of resolution. This is, of course, a familiar idea in Computer Science, where the notion of resolution, which we handled in a very crude manner here, has been considerably refined, e.g., in the technique of *multiscale representation* in which by convolving an original image with Gaussian kernels of different variances we obtain a series of images at different resolution levels (see [1, Ch.7]).

Unfortunately, the Fregean construction cannot be achieved in this instance, because the “same shape” relation on physical objects, as defined above, is not an equivalence relation. It is perfectly possible to have three objects A, B, and C, such that, at some resolution  $h$ , A has the same shape as B and B has the same shape as C, but A does not have the same shape as C. This is essentially because the “same shape” relation, as here defined, is not capturing a notion of *identical* shape so much as a notion of *indiscernible* (at resolution  $h$ ) shape; and it is a familiar fact that unlike identity relations, indiscernibility relations are not transitive. The crucial implication of this, for us, is that there is no coherent

---

<sup>5</sup>This is somewhat oversimplified since in practice the resolution of our observations will not harmonise with the levels of resolution available in the system of units used for recording them—e.g., given resolving power (for lengths) of 0.03mm, recording to 1 decimal place is too coarse and recording to two decimal places is too fine. For an attempt to deal with this issue in an ontological framework, see [10].

notion of “exact shape” for a physical object, only that of objects being more or less approximately the same shape as other objects.

#### 4.3. Similarity between physical and geometrical objects

It is often said that the shapes of physical objects can also approximate to the shapes of ideal geometrical objects (compare [11]). Thus there are many approximate spheres in the physical world, but no geometrically exact spheres, the exact sphere being an inhabitant of mathematical, not physical space. This is true so far as it goes, but we can explain what is being said here more carefully using resolution-based shape comparison. An explanation is needed since on the face of it there is something paradoxical about comparing something physical with an abstract mathematical construction: the two seem to belong to such entirely different realms that any notion of comparison ought to be out of the question.

If we are to compare the shapes of, say, a sheet  $P$  of A4 paper and a certain rectangle  $R$  defined in  $\mathbb{R}^2$ , then we need some way of matching up points in the former with points in the latter. There are, at least on the orthodox view, uncountably many points in  $R$ , but there does not seem to be any meaningful sense in which we can attribute uncountably many points to  $P$ . There is already something dubious about the notion of attributing *infinitely* many points to  $P$ , since as noted earlier,  $P$  can only be observed at all at some finite level of resolution, and at any such level only finitely many points can be distinguished within it. One might, of course, entertain the notion that, if there is no limit to how fine the resolution level can be, then there is no limit to how many points we can discern in  $P$ , so that the number of points in  $P$ , if not actually infinite, is at least *potentially* infinite. But it is far from clear that, in the physical world, resolution *could* be made indefinitely fine. For example, in the state of our current understanding of physics, the Planck length (approximately  $1.6 \times 10^{-35}$  m) is believed to provide a lower bound for the resolution of any *possible* technique of measurement.

But there is another problem we need to face, which is that while the physical piece of paper  $P$  does have an actual (albeit indeterminate) size, the mathematical rectangle  $R$  does not. How wide is the rectangle whose corners are at the points  $(0, 0)$ ,  $(0, 1)$ ,  $(\sqrt{2}, 0)$ ,  $(\sqrt{2}, 1)$ ? That’s easy: it’s  $\sqrt{2}$ ! But  $\sqrt{2}$  is a number, not a length. Well, then, it’s  $\sqrt{2}$  units. But what is a unit? How does a unit compare with a millimetre or an inch? It is a meaningless question: Objects in the mathematical world do not have actual sizes that can be compared directly to those of physical objects.

In fact this is not as serious a problem as it may seem at first sight, since in assessing geometric similarity we are only interested in relative length, not absolute length. The scale factor  $\kappa$  can take care of differences in absolute length, so long as the relative lengths remain unchanged. Thus in the ideal geometric rectangle, considered as a set of points in  $\mathbb{R}^2$ , we can follow the usual practice of recording the width as  $\sqrt{2}$  units and the height as 1 unit, even though “unit” does not designate any actual physical length, since what matters for our purposes is only the ratio between the width and the height. And clearly all geometric rectangles with sides in the ratio  $1 : \sqrt{2}$  are geometrically similar to one another.

What, then, does it mean to say that a piece of paper, considered at resolution  $h$ , has the shape of a rectangle with sides in a given ratio? We cannot use the

definition of strict geometric similarity for geometric figures here, nor can we use the definition of “same shape at resolution  $h$ ” for physical objects. Instead, we modify the criterion by combining elements from the two definitions as follows:

**Definition of a physical object having the “same shape” as a geometrical object.** At resolution  $h$ , a physical object  $P$  has the same shape as a geometrical object  $Q$  if there is an injective mapping  $\phi$  from the set of points  $S_h(P)$  discernible in  $P$  at resolution  $h$  into the set of points in  $Q$  such that, for some constant  $\kappa > 0$ :

1.  $\forall x, y \in S_h(P). \Delta(\phi(x), \phi(y)) = \kappa \Delta_h(x, y)$
2.  $\forall x \in Q. \exists y \in S_h(P). \Delta(x, \phi(y)) \leq \frac{1}{2} \kappa h.$

The point here is that the objects in the pure geometric world, being given to us by thought rather than by observation, can be specified with infinite precision: in particular, the distances between points can be arbitrary real numbers and are not constrained to being multiples of some minimal discernible distance.

## 5. Intrinsic vs Embedded Shape

Up to this point in our discussion, there has been an implicit assumption that the shape of an *object* can be identified with the shape of the portion of space, or *region*, which is occupied by that object. This fits in with a general presumption that the spatial properties of objects are nothing other than the properties of the portions of space they occupy. For many purposes this is not an unreasonable presumption, and has the advantage that all such properties can then be handled purely within a theory of space itself, without our having to worry about other physical properties of the objects.

In the case of shape, this does not always accord with our everyday ways of thinking. A rectangular sheet of paper, for example, is still, essentially, a rectangular piece of paper if it is folded along a diagonal, or screwed into a ball. A single long strand of wool retains this character whether it is coiled into in a skein or knitted into a scarf. A tall, thin person is still a tall thin person whatever posture he adopts; and more generally, the shape of a human body might be understood to be “something which is invariant across all the various postures that the body is capable of assuming” [3, p.201]. In such cases, we cannot identify the shape of the object with the shape of the space it occupies, since the shape, understood in this more general sense, may remain unchanged even as the object occupies a succession of differently-shaped spatial regions.

Let us distinguish between, on the one hand, the *intrinsic shape* of an object—which is the more general, abstract notion of shape described in the previous paragraph—and, on the other, its *embedded shape*, which is the shape of the region of space that it currently occupies. The idea is that the embedded shape of an object may change while its intrinsic shape remains constant. If we consider a square sheet of paper, and all the myriad origami figures which it can be folded into without tearing, we can say that across this range of figures the paper retains its intrinsic shape (i.e., square) but assumes different embedded shapes.

Like embedded shape, the notion of intrinsic shape is logically dependent on a prior notion of *same shape*: same embedded shape or same intrinsic shape

respectively. Both of these notions can be defined in terms of the distances between points in the respective objects. The difference is that whereas with “same embedded shape” the distances are measured with reference to the space within which the objects are embedded, with “same intrinsic shape” they must be measured within the object itself. Given a physical object  $P$ , I define the *P-intrinsic distance* between two points  $x$  and  $y$  in  $P$ , written  $\Delta_P(x, y)$ , as the length of the shortest path between  $x$  and  $y$  which lies wholly within  $P$ . I contrast this with the *embedded distance*  $\Delta(x, y)$  used previously. Note that, for points within a convex object, the intrinsic and embedded distances are the same.

We can now give a rough definition of “same intrinsic shape” as the existence of a bijective mapping  $\phi$  between the points of  $P$  and the points of  $Q$  such that, for any pair of points  $x$  and  $x'$  in  $P$  we have  $\Delta_Q(\phi(x), \phi(x')) = \kappa \Delta_P(x, x')$ , for some constant  $\kappa \in \mathbb{R}^+$ . This is only a rough definition because, since we are here dealing with physical objects, we have to take into account the resolution of the distance measurements, just as we did in the embedded case. When a person adopts different bodily postures, although their intrinsic shape remains approximately the same, we will always find, if measuring sufficiently precisely, that there are differences resulting from muscular contractions and extensions which distort the shapes of individual parts of the body. Similarly, when a piece of paper is folded, at the fold there will be minute tears or stretches which disrupt the exact relationships between the intrinsic distances. But by measuring at a sufficiently coarse resolution these small-scale disruptions will disappear from view, so that intrinsic shape, *at that resolution level*, remains constant.

For a more exact definition, then, we need to introduce the notation  $\Delta_{P,h}(x, y)$  to mean the  $P$ -intrinsic distance between  $x$  and  $y$  at resolution  $h$ . We then have:

**Definition of “same intrinsic shape” for physical objects.** Physical objects  $P$  and  $Q$  (where  $Q$  is at least as big as  $P$ ) have the same intrinsic shape, at resolution  $h$ , if, for some constant  $\kappa \geq 1$ , the set  $S_h(P)$  of points discernible in  $P$  at resolution  $h$  can be mapped into the set  $S_h(Q)$  of such points of  $Q$  by means of an injection  $\phi$ , such that the following relations hold:

1.  $\forall x, y \in S_h(P). |\Delta_{Q,h}(\phi(x), \phi(y)) - \kappa \Delta_{P,h}(x, y)| \leq \frac{1}{2}h$
2.  $\forall x \in S_h(Q). \exists y \in S_h(P). \Delta_{Q,h}(x, \phi(y)) \leq \frac{1}{2}\kappa h$

What this definition does not tell us is how widely applicable the notion of intrinsic shape is. The only examples I have given so far concern sheets of paper, strands of wool, and human bodies, but for many objects the notion of an underlying shape which is retained even as the object occupies differently-shaped regions of space does not seem to apply. We can arrive at a rough characterisation of the kinds of object for which the notion of intrinsic shape can do useful work by considering what kinds of transformations can change the embedded shape of an object while retaining its intrinsic shape.

Rigid motions—rotations, translations, and reflections—preserve *both* intrinsic and embedded shape. For rigid bodies, therefore, intrinsic shape does not convey any information beyond embedded shape. The same applies to magnification; we can see this, at least approximately, in the case of a spherical rubber ball that is being inflated; as it gets bigger, all the inter-point distances, whether measured across the embedding space or within the material of the ball itself, expand at the same rate, thus preserving both embedded and intrinsic shape.

All these transformations preserve the topology of the object, and it is certainly true that transformations that alter the topology, such as tearing or fracturing, will also alter the intrinsic shape—thus two sheets of A4 paper lying flat on the table, one of which is intact and the other has a razor-thin slit in the middle (making it topologically a torus), exhibit the same embedded shape but different intrinsic shapes. However, topological equivalence is a much coarser relation than having the same intrinsic shape: in general stretching and compression disrupt intrinsic shape while preserving topology. As noted above, transformations such as folding or “reposturing” always involve some such disruptive transformations, but because these are small compared with the objects in which they occur, intrinsic shape can be preserved even under reasonably fine resolution.

Can we characterise exactly those types of object which typically undergo transformations of a kind that result in changes of embedded shape while preserving intrinsic shape at an appropriate level of resolution? Because notions such as “typically” and “appropriate” are inherently inexact, we will never find such an exact characterisation; but it would be interesting at any rate to find a more exact characterisation than we have at present.

## 6. Conclusion

As with number and direction, the ontological status of shape is problematic because of its dependent character: shapes do not exist “in their own right”, but only as qualities of objects. As Frege observed, for the shapes of geometrical figures characterised mathematically as subsets of  $\mathbb{R}^n$ , the relation of geometrical similarity provides a robust criterion of identity which, by establishing the notion of “same shape” as an equivalence relation, can support the notion of shapes as entities that could be included within an ontology.

By contrast, the “same shape” relation for physical objects, since it can only be defined relative to some finite resolution level, fails to be an equivalence relation, and therefore cannot provide a criterion of identity for a notion of physical shape. This casts doubt on the ontological integrity of the notion of shape, and we are left with the intransitive “same shape at resolution  $h$ ” relation as the primary vehicle for shape-attribution to physical objects. This is reflected in the fact that, in practice, when we ascribe a shape to an object or object part it is always by comparison with something else—either another object or a geometrical figure—and relative to some (often implicit) level of resolution.

As a final observation, it should be noted that since “same shape” relations are founded on the comparison of distances amongst the points within the objects to be compared, it follows that a notion of shape should be, in principle, available in any domain where some notion of “distance” is applicable.

The notion of intrinsic, as opposed to embedded shape arises as a result of reinterpreting what is meant by distance: instead of distances measured across the space in which the objects under consideration are embedded, we measure distance along paths which are constrained to lie within the objects themselves. In this case “distance” is still essentially spatial, but by extending this term to measures of separation along non-spatial dimensions we obtain metaphorical extensions to the notion of shape.



One example is in the temporal domain, where we often speak of the *profile* of a process, meaning by this its temporal “shape”—typically rendered visible in the form of the spatial shape of some graphical representation of the process in which time is the independent variable and the dependent variables measure one or more qualities whose values are modified by the process. If we consider the process itself rather than its graphical representation, we are faced with the problem that the time dimension is not commensurable with the dimensions along which the other variables are measured: thus, for example, given a scale model of a railway train, where the non-temporal dimensions are spatial, there is no determinate answer to the question how fast the model should be run to preserve the spatio-temporal “shape” of the process in order to secure maximum verisimilitude.

An interesting extension of this is the idea of the “shape” of a musical phrase. In music, there are two dimensions that can provide analogues of distance, namely time and pitch.<sup>6</sup> Since these have different measurement scales, distances along these dimensions cannot be compared with each other, and this means that we can reasonably regard, e.g., augmentations or compressions along either the time axis or the pitch axis as in some sense shape-preserving. An examination of music from many different styles will furnish numerous examples of composers exploiting the expressive potential of such phenomena.

## References

- [1] L. da Fontoura Costa and R. M. Cesar. *Shape Analysis and Classification: Theory and Practice*. CRC Press, 2001.
- [2] G. Frege. *Die Grundlagen der Arithmetik*. Wilhelm Koebner, Breslau, 1884. English translation *The Foundations of Arithmetic* by J. L. Austin, Oxford: Basil Blackwell, 1950.
- [3] A. Galton. *Qualitative Spatial Change*. Oxford University Press, Oxford, 2000.
- [4] A. Galton. Operators vs arguments: The ins and outs of reification. *Synthese*, 150:415–441, 2006.
- [5] A. Galton and M. Duckham. What is the region occupied by a set of points? In M. Raubal, H. Niler, A. Frank, and M. Goodchild, editors, *Geographic Information Science: Fourth International Conference GIScience 2006*, pages 81–98. Springer, 2006.
- [6] P. Gärdenfors. *Conceptual Spaces: the geometry of thought*. MIT Press, Cambridge MA and London UK, 2000.
- [7] IFOMIS. Basic formal ontology (BFO). <http://www.ifomis.org/bfo>, Accessed 21/12/2012, 2009.
- [8] E. J. Lowe. *The Four Category Ontology. A Metaphysical Foundation for Natural Science*. Oxford University Press, Oxford, 2006.
- [9] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari. WonderWeb deliverable D18. Technical report, Laboratory for Applied Ontology, ISTC-CNR, Trento, Italy, 2003.
- [10] Florian Probst. Observations, measurements and semantic reference spaces. *Applied Ontology*, 3:63–89, 2008.
- [11] R. Rovetto. The shape of shapes: An ontological exploration. In J. Hastings, O. Kutz, M. Bhatt, and S. Borgo, editors, *SHAPES 1.0: The Shape of Things 2011 (Proceedings of the First Interdisciplinary Workshop)*, volume 812. CEUR Workshop Proceedings, 2011.
- [12] L. Schneider. The ontological square and its logic. In C. Eschenbach and M. Grüninger, editors, *Formal Ontology in Information Systems: Proceedings of the Fifth International Conference (FOIS 2008)*, pages 36–48. IOS Press, 2008.
- [13] A. Stroll. *Surfaces*. University of Minnesota Press, Minneapolis, 1988.

---

<sup>6</sup>Leaving aside loudness here as a third candidate.

# Lines, Shapes, and Meaning

Barbara TVERSKY<sup>1</sup>

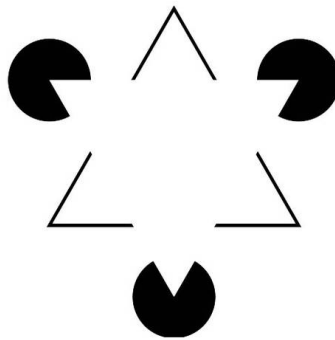
*Columbia Teachers College and Stanford University*

**Abstract.** Lines are mysterious. They are drawn by the hand, they are seen by the eye, they appear in the world. Lines are what the hands draw, what the eyes see, and what the page represents. Lines form forms. Simple regular visual/spatial forms like dots, lines, and containers, have meanings that are readily apparent in context. They are used in the service of clear communication, to self and others, notably in diagrams and gesture. They are used to organize, indeed to diagram, the world. Other, messy sketchy lines are used for exploration and discovery.

**Keywords.** diagrams, sketches, visual communication, cognitive artifacts, abstraction, creativity, design

## Introduction

We see lines everywhere. We see them even where they aren't there at all, courtesy of Kanizsa:



The eye generates them by connecting the dots, creating the continuous contours that allow us to discern objects in the camouflage of shadows and occlusion. The natural world provides them, the plane of the earth as it meets the sky and the perpendiculars of the things that grow from it. The hand draws them on paper and in the air to represent other things, concrete and abstract. The designer constructs them, connecting buildings and towns along streets, books and dishes on shelves. We travel on them as we go from place to place. Are all these lines, real or virtual, in the world or in the mind, connected?

---

<sup>1</sup> E-mail: [btversky@stanford.edu](mailto:btversky@stanford.edu)

## Creating Meaning: Orderly Lines

Lines on paper. Let's begin with the abstract, with lines that represent, lines that flow from the pen held by the hand. The simplest line is just a dot, a point, hardly there at all. Next, a line, a point extended in one dimension. After that, lines with bends and turns, zig-zags and swirls, arrows and writing. Lines that catch their tails, that close, form shapes, simple forms like circles and squares and more complex forms like contours of animals or trees. Lines create meaning by creating recognizable forms like ducks and rabbits, wives and mothers-in-law. They also create meaning by forming shadows, words, sketches, and diagrams.

Diagrams are constructed from lines. Typically, diagrams are meant to simplify a more complex state of affairs to inform or influence or instruct. A frequent kind of diagram is based in forms that bear physical resemblance to what they are intended to represent. Maps are a prototype for this kind of diagram, but also other diagrams, for example, those designed to show the crucial parts of the heart or an engine, to show how the heart pumps or an engine works, to show how to assemble a heart or an engine. Yet, such depictions in diagrams are not typically designed to be realistic renderings but rather diagrammatic renderings; that is, they may omit or exaggerate or reorganize physical appearance and structure, and they may add information, verbal, symbolic, and visual in the service of informing or instructing. Part of what makes them diagrams and not simple depictions is the addition of simple forms, notably, dots, lines, and arrows that organize, label, integrate, explain, extend, and otherwise add to the depictive information about appearances of parts and wholes. Other frequent kinds of diagrams, notably charts and graphs, don't generally bear physical similarities to what they are meant to represent, relying primarily on simple visual forms and spatial relations, and of course language and symbols, to show data and relationships.

Commonly, these simple visual forms, dots, lines, and blobs, and spatial relations, center, up/down, left/right, carry meanings that are readily understood in context. The meanings seem to derive from their geometric and gestalt properties (e. g., Tversky, 2011 a, b). Consider networks, arguably the simplest diagram, in its most rudimentary form, two dots and a line. Networks are constructed from dots and lines; the dots are nodes, the lines, edges. Abstractly, the dots are idea or entities, the lines, the relations between them. Variations of networks are used to represent myriad concepts: the network of roads on the ground, or airline routes in the air, of computers on the net, of concepts in a semantic net. Varieties of networks can represent the phylogenetic tree, a family tree, a corporate organization, a set of social connections, the transmission of ideas over time, data points along a dimension. What is shared is that the nodes represent entities and the lines represent the links among them. Why are dots understood as concepts, places, people, computers, roles, functions and more and lines understood as relations among them? River boat navigators who do not read and have not seen maps when asked to sketch a map of their travels draw settlements on the river as dots and the river routes that connect them as lines, pearls on a string (Woodward and Lewis, 1998), just as typical sketch maps (e.g., Tversky and Lee, 1998; 1999). Notably, the links form straight lines, despite the geographic irregularities of the river and the dots are the same size despite variations in size and extent. That information, the exact forms of the river and the communities, is not relevant for showing the route. What matters is the ordering and connection of the locations. For these ends, the communities are conceived of as points of zero dimensionality, the river as a line of a

single dimension. Language (at least English) makes a similar distinction (Talmy, 1983): the journey began *at* the Capitol and *continued* to the White House; any two locations could be substituted for those venerable ones.

Journeys can be one-way, relationships can be asymmetric. For that, asymmetric lines are needed, and arrows serve that need nicely. Arrows are asymmetric lines, and have a basis in the experienced world: arrows shot from a bow fly in the direction indicated; arrows form in the mud in the direction of erosion. Arrows change the meaning of lines. When asked to describe a diagram of a mechanical system without arrows, people give structural descriptions; they list the parts and their spatial relations. When asked to describe a diagram of the same systems with arrows, people provide functional descriptions; they give a step-by-step description of behavior, process, causality of the system (Heiser and Tversky, 2006).

Often a broader perspective is needed, two dimensions, not a single one. Not *at* or *to*, but *in*. Not simple dots, but containers. The cells of tables or the bars in graphs or circles on graphs or maps contain and represent many entities, those that hold the properties defined by the cells: the cars manufactured in India in 1998, the number of spectators at soccer games in Brazil in 2004. The variations that characterize each of the cars and each of the spectators is irrelevant; all that matters is their numbers; similarly, their exact locations and the exact time of the year are irrelevant. They are represented as featureless numbers by simple containers, cells, bars, and circles. This third distinction, container, is also made in language, signified by the preposition *in* (Talmy, 1983).

Why these regular, almost perfect forms; why not blobs of uncertain shape? After all, we don't know the exact shapes, so why is a perfect idealized shape used to represent a shape that might actually be known? Just as "red" represents a range of shades of red including magenta whereas "magenta" is a more specific shade of red, a purplish red (e. g., Rosch, 1978) and "noon" represents a range of times from about 11:50 to 12:15 whereas 12:02 is a specific minute—but not a specific second within that minute—circles and rectangles and squares represent a range of shapes. Just as *red* and *noon* are prototypic linguistic categories that contain and represent a set of values, dots, circles, and lines are prototypic spatial categories that contain and represent a set of values. Just as on average, the myriad shades of red contained in the category would average to the prototypic red, the myriad shapes that are contained in the visual categories points, lines, and containers would average to the prototypic values.

Again like linguistic categories, these spatial categories allow inferences, and the convergences of those inferences are evidence for their meanings, converging with their geometric and gestalt properties. People interpret bars in graphs as discrete comparisons, lines as trends; similarly, they produce bars to represent discrete comparisons and lines to represent trends (Zacks and Tversky, 1999). People invent boxes to represent categorical relations and lines of various thicknesses to represent continuous ones; their inferences follow the same patterns (Tversky, Corter, Yu, Mason, and Nickerson, 2012). Graphs of people, place, and time that connect the people with lines encourage inferences about movement of people in time whereas tables of the same information encourage a broader range and number of inferences (Kessell and Tversky, 2011).

Diagrams, maps, charts, and graphs select relevant information and omit irrelevant information. They may exaggerate, even distort, the relevant information for emphasis, for readability. Overall, their goal is clear and unambiguous communication. These simple spatial forms, and others like them, are their units of meaning, like morphemes,

of a visual/spatial vocabulary for communication. They can be combined systematically, like rules of syntax, to form genres of diagrams, sketch maps, circuit diagrams, architectural plans.

### **Finding Meaning: Messy Lines**

Sketches are a very different case of lines on paper. They use lines, the lines may form shapes, but they are typically uncertain, tentative, unclear. They represent ideas, but amorphous ideas, ideas that are not yet fully formed. Their very lack of certainty creates ambiguity, allowing many interpretations, not a single clear one. Perhaps for that reason, messy lines in sketches encourage exploration and discovery, they promote new ideas. Messy sketches are used productively by artists, designers, and architects, especially experienced ones, to explore a domain with impunity, and to generate new ideas (Goldschmidt, 1994; Kantrowitz, in preparation; Schon, 1983; Suwa and Tversky, 1996; Suwa and Tversky, 2001; Suwa and Tversky, 2003; Suwa, Tversky, Gero, and Purcell, 2001; Tversky and Chou, 2010, Tversky and Suwa, 2008). Experienced architects, artists, and designers use their sketches in deliberate ways to get new ideas, a process we have called *Constructive Perception* (Suwa and Tversky, 2003; Tversky and Suwa, 2008). One common strategy adopted by experienced architects and designers is to deliberately reconfigure their sketches, to reorganize the parts and wholes. The reconfigured perceptual array suggests new objects, encourages new interpretations, even ah-ha experiences. Reconfiguration is one of several strategies that promote constructive perception in the service of new ideas. Interspersing other tasks that expose thinkers to other perceptual and conceptual ideas help (Tversky and Chou, 2010). Prompted hints help, especially top-down hints, thinking of other domains (Tversky and Chou, in progress). Talent helps, two talents actually, a perceptual talent, measured by the ability to see elemental complex figures embedded in more complex ones (Gottschaldt, 1926) and a cognitive talent, the ability to make remote associations among ideas (Mednick and Mednick, 1967) (Suwa and Tversky, 2003). These talents capture the two aspects of constructive perception, perceptual reconfiguration and conceptual interpretations.

Now the caveats. Of course, messy lines aren't always ambiguous, and ambiguity isn't always productive. Similarly, orderly lines do not always convey a clear message.

### **Lines in the World, Lines in the Brain, Lines on the Page**

Lines are what the hand draws and what the eye sees, a magical convergence. The world on the retina is pixels, dots, representing light of varying brightness and hue. The brain connects the dots, forming lines and shapes that constitute the uncountable number of things we recognize. Designers sketch lines, first messy tentative ones that allow interpretation and reinterpretation, evolving into orderly forceful lines that can convey myriad ideas with clarity.

**Acknowledgements.** The following grants facilitated the research and/or preparation of the manuscript: National Science Foundation HHC 0905417, IIS-0725223, IIS-0855995, and REC 0440103, the Stanford Regional Visualization and Analysis Center, and Office of Naval Research N00014-PP-1-O649, N000140110717, and N000140210534.

## References

- [1] Goldschmidt, G. (1994). On visual design thinking: the vis kids of architecture, *Design Studies*, 15, 158-174.
- [2] Gottschaldt, K. (1926). Über den Einfluss der Erfahrung auf die Wahrnehmung von Figuren, I, *Psychologische Forschung*, 8, 261-317.
- [3] Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science*, 30, 581-592.
- [4] Howard-Jones, P. A. (1998). The variation of ideational productivity over short timescales and the influence of an instructional strategy to defocus attention, *Proceedings of Twentieth Annual Meeting of the Cognitive Science Society* (pp. 496-501), Hillsdale, NJ: Lawrence Erlbaum Associates.
- [5] Kantrowitz, A. (in progress). Dissertation.
- [6] Kessell, A. M. and Tversky, B. (2011). Visualizing space, time, and agents: Production, performance, and preference. *Cognitive Processing*, 12, 43-52.
- [7] Mednick, S.A., & Mednick, M.T. (1967). *Examiner's manual: Remote Associates Test*. Boston: Houghton Mifflin.
- [8] Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum Associates.
- [9] Schon, D. A. (1983). *The reflective practitioner*. New York, NY: Basic Books.
- [10] Suwa, M. and Tversky, B. (1996). What architects see in their sketches: Implications for design tools. *Human factors in computing systems: Conference companion* (pp. 191-192). NY: ACM.
- [11] Suwa, M., & Tversky, B. (2001). Constructive perception in design. In J. S. Gero & M. L. Maher (Eds.) *Computational and cognitive models of creative design V*. Pp.227-239. Sydney: Key Centre of Design Computing and Cognition.
- [12] Suwa, M. and Tversky, B. (2003). Constructive perception: A skill for coordinating perception and conception. In *Proceedings of the Cognitive Science Society Meetings*.
- [13] Suwa, M., Tversky, B., Gero, J., and Purcell, T. (2001). Seeing into sketches: Regrouping parts encourages new interpretations. In J. S. Gero, B. Tversky, and T. Purcell (Editors). *Visual and spatial reasoning in design*. Pp. 207-219. Sydney, Australia: Key Centre of Design Computing and Cognition.
- [14] Talmy, L. (1983). How language structures space. In: H. L. Pick, Jr. & L. P. Acredolo (Editors). *Spatial orientation: Theory, research and application*. Pp. 225-282. New York: Plenum.
- [15] Tversky, B. (2011a). Spatial thought, social thought. In T. Schubert and A. Maass (Editors). *Spatial schemas in social thought*. Berlin: Mouton de Gruyter.
- [16] Tversky, B. (2011b). Visualizations of thought. *Topics in Cognitive Science*, 3, 499-535.
- [17] Tversky, B., Corter, J.E., Yu, L., Mason, D. L., and Nickerson, J. V. (2012). Representing category and continuum: Visualizing thought. In P. Rodgers, P. Cox and B. Plimmer (Editors). *Diagrammatic representation and inference*. Pp. 23-34. Berlin: Springer.
- [18] Tversky, B. and Chou, J. Y. (2010). Creativity: Depth and breadth. In Y. Nagai (Editor). *Design creativity*. Dordrecht, Netherlands: Springer.
- [19] Tversky, B., & Lee, P. U. (1998). How space structures language. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial cognition: An interdisciplinary approach to representation and processing of spatial knowledge* (pp. 157-175). Berlin: Springer-Verlag.
- [20] Tversky, B., & Lee, P. U. (1999). Pictorial and verbal tools for conveying routes. In C. Freksa & D. M. Mark (Eds.), *Spatial information theory: Cognitive and computational foundations of geographic information science* (pp. 51-64). Berlin: Springer.
- [21] Tversky, B. and Suwa, M. (2009). Thinking with sketches. In A. Markman (Editor), *Tools for innovation*. Oxford: Oxford University Press.
- [22] Woodward, D., & Lewis, G. M. (1998). *History of cartography. Vol. 2. Book 3: Cartography in the traditional Africa, America, Arctic, Australian, and Pacific societies*. Chicago: University of Chicago Press.
- [23] Zacks, J., & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory and Cognition*, 27, 1073-1079.



## Contributed Talks





# Sketch Learning by Analogy

Ulf KRUMNACK<sup>a</sup>, Angela SCHWERING<sup>b</sup>, Kai-Uwe KÜHNBERGER<sup>a</sup>,  
Helmar GUST<sup>a</sup>, Ahmed ABDEL-FATTAH<sup>a,1</sup>, Tarek BESOLD<sup>a</sup>,  
Martin SCHMIDT<sup>a</sup>, and Stefan SCHNEIDER<sup>a</sup>

<sup>a</sup>*Institute of Cognitive Science, University of Osnabrück, Germany*

<sup>b</sup>*Institute for Geoinformatics, University of Münster, Germany*

**Abstract.** Sketches are shapes that represent objects, scenes, or ideas by depicting relevant parts and their spatial arrangements. While humans are quite efficient in understanding and using sketch drawings, those are largely inaccessible to computers. We argue that this is due to a specific shape based representation by humans and hence the use of cognitively inspired representation and reasoning techniques could lead to more proficient sketch processing. We also propose a three-level architecture for sketch learning and recognition that builds on concepts from cognitive science, especially from analogy research, to map and generalize sketches.

**Keywords.** Sketch, Shape, Learning, Analogy

## 1. Introduction

Sketches can be considered as an intermediate level of abstraction between raw sub-symbolic streams of sensory input on the one side and icons on the other. In contrast to a drawing, a sketch only captures the conceptually relevant parts of the displayed object or situation as well as the spatial relations between these parts, making their treatment substantially different from classical image processing. The pertinence of sketches for future information technology applications and services can hardly be overestimated. Especially the spread of tablet computers and devices equipped with touch screens paves the way for new human computer interfaces, in which sketches can play an essential role. Future applications can be search services for large knowledge bases utilizing input sketches, support services in software systems for shortening the path through complex menus, automatic sketch generation for manuals and assembly instructions, a bridging approach between computer vision and conceptual reasoning, or creative usage of sketches in e-learning contexts.

In this paper, we present ideas on modeling the human ability to operate with sketches. We focus on a competence model for recognition, classification, memorization and retrieval of sketches guided by cognitive principles. In a first step, the envisaged system acquires basic knowledge on how to sketch a given

---

<sup>1</sup>Authors are listed in alphabetical order.

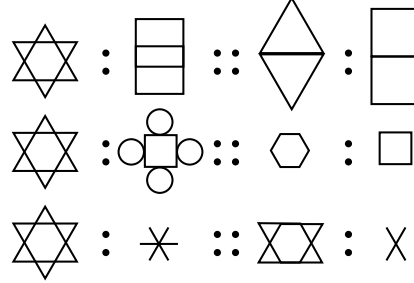
object. The essential and optional components as well as their spatial arrangement are learned by comparing different sketches of the same type of object provided to the system as training data. In the next step, after elementary types have been learned in this bootstrapping process, the system will generate more abstract categories by cross-type comparison, establishing a hierarchical index of sketch schemata and shapes. This index will then support the recognition capacity: new sketches will be compared to the abstract descriptions in the sketch database to find structurally matching sketches in memory. We argue in favor of a symbolic approach because the structure of a sketch can be captured explicitly in such a representation, and changes in the conceptualization can be performed by automatic inference techniques.

The paper is structured as follows. We start with discussing requirements for a representation language for sketches in section 2. The description of the proposed system is given in section 3, which constitutes the main part of this paper. We then provide links to related work in section 4, before concluding with some remarks and future work in section 5.

## 2. Sketch Representation and Re-representation

Sketches are assumed to be given as a collection of dots and lines, possibly annotated with an order of drawing. Multiple relational representations can thus be constructed based on psychological principles, which take into account that human cognition of spatial environments is qualitative in nature. Humans do not perceive absolute locations and quantitative relations between spatial objects, but rather relative locations and qualitative relations [1,2,3,4]. By observing a geometric figure, the unstructured information is transformed into a structured representation of coherent shapes and patterns [5,6]. Perception tends to follow a set of Gestalt principles: stimuli are experienced as a possibly good Gestalt, i.e. as regular, simplistic, ordered, and symmetrical as possible. Gestalt psychology argues that human perception is holistic: instead of collecting every single element of a spatial object and afterwards composing all parts into one integrated picture, people experience things as an integral, meaningful whole. The whole contains an internal structure described by relationships among the individual elements.

We argue that qualitative spatial relations play a major role during sketch recognition and hence sketches should be described on a qualitative level by a symbolic language. The spatial representation language has to meet two major requirements: it must describe all elements of a spatial object with respect to the aspects relevant in human perception, and it must also describe the spatial characteristics that are important in recognizing spatial objects. To reflect human perception, the language must comprise significant perceptual vocabulary to specify visual structures. The geometry in a sketch, i.e. of its elements and their spatial relations, has to be represented in a way that allows for cognitively plausible reasoning. The language can be based on psychological theories for perception and pattern recognition, such as Gestalt Theory [7,8,5,6], Marr’s theory of vision [9] and Biedermann’s Geons [10], and on research specifically directed towards the sketch mapping task such as the CogSketch [11] approach.



**Figure 1.** Visual ambiguity exemplified by different representations of a sketch relative to a given context in a proportional analogy, according to [12].

The potential ambiguity of sketches, e.g. caused by different groupings of elements or different interpretations, is an essential point to be considered. Indurkha [12] has demonstrated the effects of visual ambiguity in proportional geometric analogies and has argued for a mechanism that can change representations. The importance of re-representation is exemplified in Figure 1, where structural commonalities between representations can be detected only if suitable representations for the geometric figure are available. The Star of David in the top row of Figure 1 should be represented as two overlapping triangles, whereas the one in the middle row should be represented as six triangles plus a central hexagon, and that in the bottom row should be represented as three overlapping rhombuses. Re-representation in this case means changing from one of these representations to another one which suits better to the given problem.

Re-representation, in the domain of sketches, means spatial re-organization and re-structuring of the elements within a spatial object, and can be formalized as a deduction task: from a given description of a sketch an alternative description has to be derived, that represents the same visual scene. It therefore requires spatial reasoning capabilities and existing qualitative spatial reasoners can be used to support this task (such as the SparQ toolbox [13] or General Qualitative Reasoner (GQR) [14]). Furthermore, to reflect human strategies of re-representation, appropriate heuristics are needed to guide the re-representation process.

### 3. A System for Analogy-Based Sketch Learning

Human learning is not a one-step action but a continuous, incremental process of acquiring new and revising old knowledge, where knowledge is learned at different levels of abstraction. Such observations about human learning motivate us to develop a three-level architecture for learning perceptual categories based on sketches. Perceptual categories in this context refer to structured representations of graphical elements that are common to a class of sketch drawings, represented as structured descriptions with respect to relevant topological, directional, and geometrical properties. The two main mechanisms for learning are learning via transfer and learning by abstraction. The former refers to the transfer of facts from the source to the target domain, while the latter denotes the generalization process that is essential to derive abstract concept definitions. Existing classical

learning approaches usually require large sets of data samples to create generalizations, though humans can already generalize over a small set of samples.

Our proposed system applies analogical comparison to discover structural commonalities and combines them with inductive refinement to extract the essential characteristics defining a perceptual category. Analogy-making, as a non-standard reasoning technique, is combined with classical deductive and inductive reasoning to compare different sketch drawings for commonalities and generalize a common underlying perceptual category. For all tasks involving comparison of sketches, analogical mapping is used to align two stimuli based on structural similarities. Such a mapping is essentially shape based, i.e. it is performed on visual descriptions only, and does not rely on functional, intensional, or usage-based information. There are two central requirements that need to be realized. The system needs, first, to be able to incrementally add newly learned categories, and, secondly, to be adaptive in the sense that a computed generalization is modifiable if new stimuli require a relaxation of the imposed constraints. Knowledge learned from training examples can be used to recognize and classify new sketches.

The model presented in this section is inspired by [15], where first ideas for an incremental learning theory were proposed. In that paper, we used a multi-layered model based on analogies to explain how abstract physical principles such as the law of energy conservation and the concept of an equilibrium of forces can be learned. These ideas are revived here and applied to the domain of sketches yielding a three-level architecture. The first level refers to the computation of analogical generalizations between a pair of sketches (section 3.1). The second level is the inductive refinement of the computed generalizations based on a re-representation process that adapts representations to make it compatible to further sketches (section 3.2). The third level focuses on learning through a revision process when comparing abstract generalizations to new domains (section 3.3). Finally, we discuss how the acquired knowledge can be used for sketch recognition (section 3.4).

### 3.1. Level 1: Analogical Generalization

At the lowest level, two sketches are taken as input, and an analogy between them is computed based on structural commonalities (cf. Figure 2). The relational structure of the description of the sketches is thus crucial. The analogical mapping may be partial, i.e. it allows parts of one sketch that have no counter-parts in

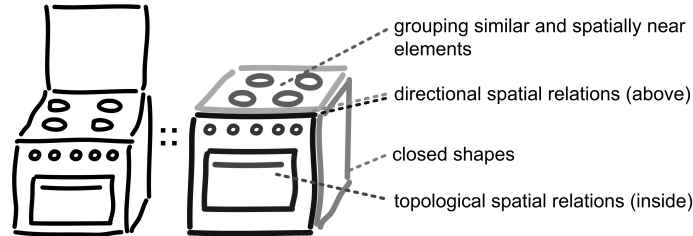


Figure 2. A flat description of a sketch is mapped to a structural representation

the other sketch. The mapping will give rise to a generalization, i.e. an abstract description of the common parts of both sketches.

Heuristic-driven theory projection (HDTP) is a logic-based framework for analogy making, presented in [16], where domains are described by logical theories and are represented by a finite set of axioms. An analogy is established by mapping axioms of two domains, based on a generalization computed via anti-unification (cf. [17]). HDTP allows re-representation of input domains: If the axiomatizations provided for the domains do not exhibit sufficient common structure to establish a good analogy, formulas from the domain theory, which can be derived from the axioms by logical deduction, are considered for mapping (cf. [18]).

The framework uses a set of heuristics to compute an analogical mapping that can be adapted to fit the special needs in the sketch domain. Essential complexity measures and heuristics are applied on different levels to guide the alignment process and to evaluate possible mappings in the sketch mapping scenario. Heuristics are used to (1) determine the order in which axioms are selected and included in the mapping process: psychologically motivated (and syntactic) heuristics can prove useful, where perceptually significant elements in human perception are likely to influence the analogy-making process more than non-significant elements (axioms should be selected therefore in the order of perceptual significance); (2) guide the re-representation: heuristics should reflect human strategies of re-representation, and the spatial language, particularly the re-representation rules, influences the development of the heuristics; and (3) determine when an analogy contains sufficient analogous structures such that a new sketch stimulus can be classified as a certain object. The approach has to bridge the gap between largest possible mappings – the more analogical structures are identified, the better the analogy – and differences in the sketches that should not be part of the analogy.

### 3.2. Level 2: Inductive Refinement

Inductive refinement is motivated by transferring ideas of concept formation to perceptual category learning. By comparing different sketches of objects, which should fall under the same category, the system should be able to construct a description of this category in terms of the relevant visual features. The inductive refinement proposed here combines a generalization of classified sketches as well as a clustering of subsets of the classified objects.

Figure 3 illustrates an example: four sketches of stoves are compared. All of them have a cubic shape and share significant elements of stoves such as hot-

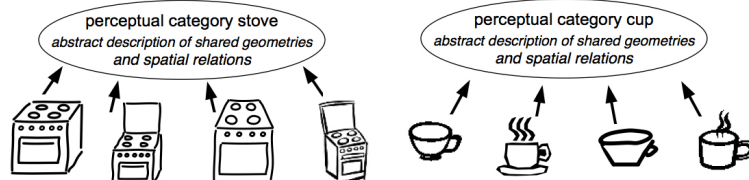
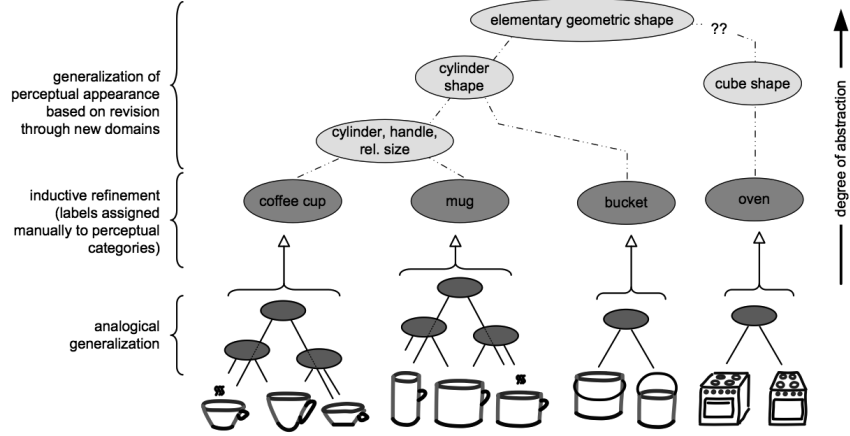


Figure 3. A structural comparison of sketches reveals commonalities that all sketches share.



**Figure 4.** Hierarchical structure of categories learned from sketches.

plates and temperature regulators. Given a pair of sketches, the first level of the proposed system detects the analogous structure and constructs a generalization containing the commonalities as a by-product. This generalization represents the first step towards the perceptual category stove at an abstract level. Iterating this process with additional input stimuli and computing generalizations of already computed generalization candidates will elaborate this category. More generally, provided a set of sketches is given, the exemplified brute force approach would be to compute for each pair of sketches a generalization. These generalizations function as candidates for new perceptual categories, and can be ordered according to their generalization complexity (e.g. substitution lengths in HDTP: The smaller the substitution lengths in the anti-unification process, the more plausible it is to assume that the two input sketches belong to the generated candidate for a perceptual category). The ordered set of candidate generalizations can be used for further structural comparison via anti-unification in order to find commonalities between more than two sketches. Applying clustering techniques may possibly identify optional elements of sketches that appear in many but not all objects (e.g. water vapor over the cups in Figure 3).

### 3.3. Level 3: Creating a Perceptual Category Hierarchy

Analogies are not only iteratively applied among instances of the same category (drawings of cups), but also between sketch drawings of different categories such as cups, mugs, buckets etc., so that a hierarchy of perceptual categories is attempted to be built (cf. Figure 4). Generated perceptual categories from Level 2 will constitute the leaves of the hierarchy. By analogical comparison of pairs of perceptual categories, generalizations are computed that can represent candidates of new, more abstract perceptual categories. These candidates can be ordered according to the complexity of the underlying analogical mapping and only those candidates constitute new categories that are maximally similar to each other. The generalizations successively reach an abstraction level such that the highest level of generalizations contains elementary geometric shapes.

### 3.4. The Recognition Task

The recognition task refers to the problem of determining whether a given sketch corresponds to an object from the system’s knowledge base. It can also be treated as an analogy problem, in which the source domain consists of the system’s knowledge on how to sketch a certain object, and has to be mapped to the unstructured graphical input (target) presented to the system as a flat collection of lines and dots. The structural commonality between the flat representation of the target and the structured representation of the source is initially not obvious. To successfully classify a new stimulus, an analogous structure has to be created for the target stimulus. During the analogy-based mapping process the target must be re-represented such that common structures may become visible.

The hierarchical memory structure built by the system (cf. Figure 4) is used as a starting point for the retrieval. The search algorithm will try to map abstract categories from that hierarchy to the search item, by computing appropriate substitutions to prove that the search item is a suitable instance of that abstract category. Hence, the retrieval is organized as a top-down search: starting from the most abstract category, all sub-categories are analogically mapped to the query sketch. Good matches are those categories where the aligned elements reach maximal coverage of the stored descriptions as well as maximal coverage of the search item. Matching items are all those sketches which are classified below a suitable category in the hierarchy. Suitable categories need to exhibit a sufficiently high coverage of the search item and the category itself. The result of a retrieval process ranks all matching items according to their relevance. We suggest the following criteria to determine the degree of relevance:

1. Depth of the matching database category: The higher a matching category in the hierarchy, the more abstract it is.
2. Coverage of the analogy: We assume that the higher the coverage of the search item, the better is the match.
3. The analogical relation between the search and the database items should be a coherent and connected match. This indicates that not only single elements align, but at least a certain part of the sketch aligns coherently.

In a ranking heuristics that combines the different aspects, the coverage has to be considered with respect to the abstractness of the database category.

## 4. Related Work

The ideas presented here build on two research fields: spatial analogies and category learning with analogies. Spatial analogies have a rather long history in artificial intelligence, whereas analogy-based learning is far less developed. The first analogy system, ANALOGY [19], was dedicated to solving proportional geometric analogy problems. O’Hara & Indurkha [20,21] proposed InterAct, an algebraic analogy model for geometric proportional analogies between line drawings. Das-tani [22] developed a formal language for this analogy model to describe elements in geometric figures and compute automatically a structural, Gestalt-based rep-



resentation. Forbus et al. [11] developed a general architecture for sketch understanding, CogSketch, which is domain independent and takes freehand sketches as input [23]. Each freehand sketch drawing consists of several primitive elements called glyphs. CogSketch interprets the primitive elements via their ontological description and via their shapes, and computes spatial relations between primitive elements based on the convex hull of glyphs. Copycat is a non-deterministic analogy model for proportional analogies in the string domain [24]. Tabletop [25] is a computational program based on Copycat that was developed to detect analogous spatial arrangements in a micro-world such as a well-laid table. Like Copycat, Tabletop combines representation-building and correspondence-finding into one integrated process. Davies and colleagues examine visual analogies in architectural design. They showed in experiments [26] that humans use visuospatial representations for the analogical mapping and transfer: participants used visual and spatial knowledge, mostly the topology of objects, to align a given architectural design with an architectural design problem and construct a solution via analogical transfer. Davies et al. developed the analogy model Galatea, an implementation of the constructive adaptive visual analogy theory [27,28], to compute visuospatial analogies.

Analogy-based learning differs from the enormous number of proposed classes of learning methods and methodologies in classical artificial intelligence research, as for example, instance-based learning, exemplar-based learning, case-based learning in the area of lazy learning and version space learning, decision tree learning, inductive learning, neural learning, and probabilistic learning in the area of eager learning. Many of these approaches require a relatively large sample of examples in order to learn reasonable generalizations. Although there may be certain approaches that attempt to incorporate structure of the generalization space into the learning process, in order to facilitate learning from small training data samples – similar to analogical learning – there are significant differences between these approaches and analogy-based learning. Only a rather limited number of positive examples are required for learning due to the conceptually guided way of establishing analogical generalizations, which are the source for new knowledge. An explicit generalization is necessary to capture new categories, re-use learned knowledge, and refine knowledge over learning steps. It is worth pointing out that one can find quite often references to analogical learning [29], but no spelled-out theory of analogical learning has been proposed so far. Inductive Logic Programming (ILP) [30] and Relational Learning [31] could be mentioned as a modern probabilistic version of frameworks where structure plays an important role in the learning approach. But compared to these most prominent approaches, the computation of an analogical relation does not incorporate probabilities, nor does it require that examples are taken from the same domain. However, the computation of an analogical relation is a complex process including aspects like retrieval, transfer, re-representation, refinement etc. Closest in spirit to analogy-making, may be the approach originally proposed by Plotkin [17], who computed least general generalizations for facilitating learning.

## 5. Summary and Future Work

We have outlined ideas for a system to model sketch learning and recognition. The setup is motivated by psychological findings emphasizing that human recognition capabilities are not only data-driven, but crucially governed by cognitive mechanisms and principles such as analogical reasoning and Gestalt principles. This contrasts with most work in the context of image retrieval, which use low-level features and does not guarantee that the resulting model reflects the human competence in recognition processes, as many of the used features are possibly not accessible by humans. One of the rare exceptions is [32] who propose to view image retrieval as a knowledge representation problem, where structured objects are retrieved such that syntactic and semantic aspects play an important role.

Even though the work presented here is currently purely conceptual, we have explained in detail how the envisaged system can make use of existing technologies, especially from the field of spatial and analogical reasoning. We have argued in favour of a symbolic representation of visual scenes and have proposed to use HDTP as a framework for analogy making. For our system, HDTP has to be extended to make use of spatial reasoners, e.g. from the SparQ toolbox [13], for re-representation during the analogical mapping. A prototype implementation may be applied to a set of test sketches, allowing to compare different heuristics. A primary concern is the development of a suitable language for describing shapes and sketches. Here we can build on a plethora of existing semiformal and formal approaches, like Dastani’s languages of perception [22]. Central objectives for such a language are, that it allows for cognitively plausible representation and supports the manipulations required by our system.

## References

- [1] C. Freksa. Qualitative spatial reasoning. In *Cognitive and Linguistic Aspects of Geographic Space*, pages 361–372. Kluwer Academic Publishers, Dordrecht, 1991.
- [2] B. Kuipers. Modeling spatial knowledge. *Cognitive Science: A Multidisciplinary Journal*, 2(2):129–153, 1978.
- [3] K. Lynch. *The Image of a City*. MIT Press, Cambridge, MA, 1960.
- [4] D.R. Montello. Spatial cognition. In N.J. Smelser and P.B. Baltes, editors, *International Encyclopedia of the Social & Behavioral Sciences*, pages 14771–14775. Pergamon Press, Oxford, 2001.
- [5] M. Wertheimer. Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, 61(1):161–265, 1912.
- [6] M. Wertheimer. *Productive Thinking*. Harper & Row, New York, 1954.
- [7] K. Koffka. *Principles of Gestalt Psychology*. Harcourt, New York, 1935.
- [8] W. Köhler. *Gestalt Psychology*. Liveright, New York, 1929.
- [9] D. Marr. *Vision*. W. H. Freeman and Company, New York, 14th ed. edition, 2000.
- [10] I. Biedermann. Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94(2):115–147, 1987.
- [11] K.D. Forbus, J. Usher, A. Lovett, K. Lockwood, and J. Wetzell. CogSketch: Open-domain sketch understanding for cognitive science research and for education. In *EUROGRAPHICS Workshop on Sketch-Based Interfaces and Modeling*, 2008.
- [12] B. Indurkha. Modes of analogy. In K.P. Jantke, editor, *Analogical and Inductive Inference*, number 397 in LNAI, pages 217–230, 1989.
- [13] D. Wolter. SparQ - a spatial reasoning toolbox. In *AAAI Spring Symposium on Benchmarking of Qualitative Spatial and Temporal Reasoning Systems*, 2009.

- [14] M. Westphal, S. Wöfl, and Z. Gantner. Gqr: A fast solver for binary qualitative constraint networks. In *AAAI Spring Symposium on Benchmarking of Qualitative Spatial and Temporal Reasoning Systems*, 2009.
- [15] H. Gust, U. Krumnack, K.-U. Kühnberger, and A. Schwering. Integrating analogical and inductive learning at different levels of generalization. In *Workshop on Learning from Non- Vectorial Data (LNVD2007) at the 30th Annual German Conference on Artificial Intelligence (KI07)*, pages 46–57, 2007.
- [16] A. Schwering, U. Krumnack, K.-U. Kühnberger, and H. Gust. Syntactic principles of Heuristic-Driven Theory Projection. *Special Issue on Analogies - Integrating Cognitive Abilities. In: J. of CSR*, 10(3):251–269, 2009.
- [17] G. D. Plotkin. A note on inductive generalization. *Machine Intelligence*, 5:153–163, 1969.
- [18] U. Krumnack, H. Gust, K.-U. Kühnberger, and A. Schwering. Re-representation in a logic-based model for analogy making. In Wayne R. Wobcke and Mengjie Zhang, editors, *AI 2008: Advances in AI, 21st Australasian AI Conference (AI-08)*, volume 5360 of *LNAI*, pages 42–48. Springer, 2008.
- [19] T. G. Evans. A heuristic program to solve geometric-analogy problems. In *Proc. of the April 21-23, 1964, spring joint computer conference*, pages 327–338, 1964.
- [20] S. O’Hara. A model of the redescription process in the context of geometric proportional analogy problems. In *Proc. of the International Workshop on Analogy and Inductive Inference (AII’92)*, pages 268–293. Springer, 1992.
- [21] S. O’Hara and B. Indurkha. Incorporating (re)-interpretation in case-based reasoning. In *Topics in Case-Based Reasoning*, volume 837 of *LNCS*, pages 246–260. Springer, Berlin, Heidelberg, 1994.
- [22] M.M. Dastani. *Languages of Perception*. PhD thesis, University of Amsterdam, 1998.
- [23] K.D. Forbus, K. Lockwood, M. Klenk, E. Tomai, and J. Usher. Open-domain sketch understanding: The nuSketch approach. In *AAAI Fall Symposium on Making Pen-based Interaction Intelligent and Natural*, Washington, DC, 2004.
- [24] D. R. Hofstadter and M. Mitchell. The copycat project: A model of mental fluidity and analogy-making. In K. Holyoak and J. Barnden, editors, *Advances in Connectionist and Neural Computation Theory Vol 2: Analogical Connections*, pages 31–112. Ablex Publishing Corporation, Norwood NJ, 1994.
- [25] R. M. French. The computational modeling of analogy-making. *Trends in Cognitive Science*, 6(5):200–205, 2002.
- [26] C. Davies, A.K. Goel, and N.J. Nersessian. A computational model of visual analogies in design. *CSR*, 10(3):204–215, 2009. Special Issue on Analogies – Integrating Cognitive Abilities.
- [27] C. Davies. *Constructive adaptive visual analogy*. PhD thesis, College of Computing,, Georgia Institute of Technology: Georgia, 2004.
- [28] J. Davies and A.K. Goel. Representation issues in visual analogy. In *Proc. of the 25th Annual Conference of the CogSci Society*, 2003.
- [29] Dedre Gentner. The mechanisms of analogical learning. In S. Vosniadou and A. Ortony, editors, *Similarity and analogical reasoning*, pages 197–241. Cambridge University Press, Cambridge, 1989.
- [30] S.H. Muggleton. Inductive logic programming. *New Generation Computing*, 8(4):295–318, 1991.
- [31] L. de Raedt. *Logical and Relational Learning*. Cognitive Technologies. Springer, Berlin, Heidelberg, 2008.
- [32] Eugenio Di Sciascio, F. M. Donini, and M. Mongiello. Spatial layout representation for query-by-sketch content-based image retrieval. *Pattern Recognition Letters*, 23:1599–1612, 2002.

# The Shape of Empty Space: Human-centred cognitive foundations in computing for spatial design

Mehul BHATT<sup>1</sup> and Carl SCHULTZ

*Spatial Cognition Research Center (SFB/TR 8), University of Bremen, Germany*

**Abstract.** We propose a human-centred model for abstraction, modelling and computing in function-driven spatial design for architecture. The primitive entities of our design conception ontology and computing framework are driven by classic notions of structure, function, and affordance in design, and are directly based on the fundamental human perceptual and analytical modalities of visual and locomotive exploration of space. With an emphasis on design semantics, our model for spatial design marks a fundamental shift from contemporary modelling and computational foundations underlying engineering-centred computer aided design systems. We demonstrate the application of our model within a system for human-centred computational design analysis and simulation. We also illustrate the manner in which our design modelling and computing framework seamlessly builds on contemporary industry data modelling standards within the architecture and construction informatics communities.

**Keywords.** architectural CAAD, cognitive systems, ontologies (artificial intelligence), declarative languages, knowledge representation and reasoning, geometric and spatial representation and reasoning, computational geometry

## References

- Bhatt, M., Schultz, C., Huang, M. (2012). The Shape of Empty Space: Human-Centred Cognitive Foundations in Computing for Spatial Design. IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC) 2012, Innsbruck, Austria
- Bhatt, M., Schultz, C., Freksa, C. (2013). The ‘Space’ in Spatial Assistance Systems: Conception, Formalisation, and Computation. in Thora Tenbrink, Jan Wiener, Christophe Claramunt (editors). Representing space in cognition: Behavior, language, and formal models. Series: Explorations in Language and Space. Oxford University Press, 2013, 978-0-19-967991-1.
- Bhatt, M., Lee, J. H., Schultz, C. (2011). *CLP(QS): A Declarative Spatial Reasoning Framework*. Proceedings of the 10th International Conference on Spatial Information Theory (COSIT 11). Editors: Egenhofer, Max and Giudice, Nicholas and Moratz, Reinhard and Worboys, Michael. Vol. 6899, Pages: 210–230, Springer Berlin / Heidelberg (LNCS), Belfast, Maine. ISBN: 978-3-642-23195-7.

---

<sup>1</sup>Corresponding Author: E-mail: bhatt@informatik.uni-bremen.de



# A new Molyneux's problem: Sounds, shapes and arbitrary crossmodal correspondences

Ophelia DEROY <sup>a1</sup>, and Malika AUVRAY <sup>b</sup>

<sup>a</sup> *Centre for the Study of the Senses, Institute of Philosophy,  
University of London, UK*

<sup>b</sup> *LIMSI, CNRS, Orsay, France*

**Abstract.** Several studies in cognitive sciences have highlighted the existence of privileged and universal psychological associations between shape attributes, such as angularity, and auditory dimensions, such as pitch. These results add a new puzzle to the list of arbitrary-looking crossmodal matching tendencies whose origin is hard to explain. The puzzle is all the more general in the case of shape that the shapes-sounds correspondences have a wide set of documented effects on perception and behaviour: Sounds can for instance influence the way a certain shape is perceived (Sweeny et al., 2012). In this talk, we suggest that the study of these crossmodal correspondences can be related to the more classical cases of crossmodal transfer of shape between vision and touch documented as part of Molyneux's question, and reveal the role that movement plays as an amodal invariant in explaining the variety of multimodal associations around shape.

**Keywords.** Crossmodal correspondences, Audition, Touch; Molyneux's problem; Amodal invariants

## Introduction: A contemporary version of Molyneux's problem

How do shapes sound? The question does not seem to make sense metaphysically: Shapes are not endowed with auditory properties. In addition, similarities or differences in shapes do not directly correlate with differences in sounds, given that crucial elements such as density, size, and material properties will make similarly shaped objects sound very differently when they are similarly struck. For instance, a small dense sphere might have the same sound as a bigger and less dense cylinder when both are struck in a similar way; and the rich repertoire of drums should convince us that shape is not all that matters to determine how objects sound.

If the question 'how do shapes sound?' needs to be dismissed then, a milder version of the question might be more resistant: Supposing that shapes have sounds, what would their sound be? Surprisingly, several studies in cognitive sciences have highlighted convergent and stable responses to this question, and they have shown the existence of privileged psychological associations between shape attributes and auditory dimensions, such as pitch. When asked which of two shapes, one rounded and the other one angular, should be called 'Takete' and which one should be called

---

<sup>1</sup> Correspondence to: Ophelia Deroy, Centre for the Study of the Senses, Institute of Philosophy, University of London, Malet Street, London, UK. [ophelia.deroy@sas.ac.uk](mailto:ophelia.deroy@sas.ac.uk)

‘Maluma’, most participants answer that the angular shape should be ‘Takete’ and the rounded one, ‘Maluma’ (Köhler, 1929, 1947; see also Ramachandran & Hubbard, 2001a and figure 1).

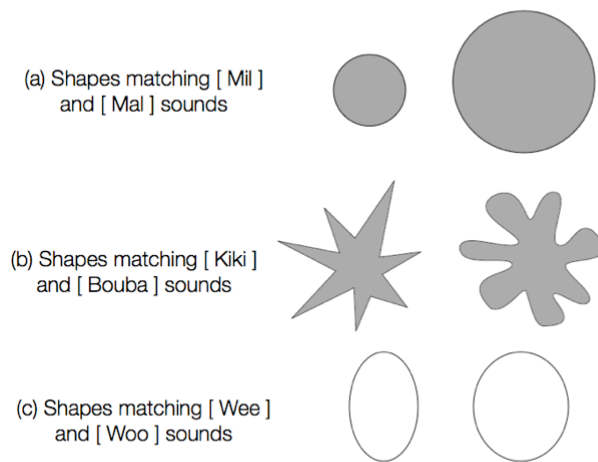


Figure 1. Three examples of crossmodal correspondences, documented between (a) sounds and size by Sapir (1929); (b) sounds and shape (angularity) by Köhler (1929, 1947) and Ramachandran & Hubbard (2001); and (c) sounds and shape (aspect ratio) by Sweeny et al. (2012).

This crossmodal association between shapes and sounds might look surprising at first, but a series of evidence shows it to be present across cultures (Bremner et al., 2013) and from an early age (i.e., four months, see Orztuck et al., 2012, see also Maurer et al., 2006, for evidence in 2 to 2,5 years old). While neurological investigation starts to unveil a specific pattern of neurological activity in the superior / intraparietal regions as well as in frontal areas corresponding to the shape-sound associations (Kovic et al., 2009; Peiffer-Smadja, 2010; see also Bien et al., 2012 for a EEG/ TMS study and Sadaghiani et al., 2009, for a fMRI study of related arbitrary audio-visual correspondences), associations between shapes and sounds is absent in individuals with damage to the angular gyrus (Ramachandran & Hubbard, 2001b), suggesting that this is a robust neuropsychological phenomenon.

What's more, shapes-sounds correspondences have recently been shown to have behavioural consequences, as the visual perception of briefly presented shapes can be affected by certain types of sounds (Sweeny et al., 2012; see also Spence & Deroy, 2012a for a discussion). Sweeny and his colleagues have indeed shown that oval shapes,

whose aspect ratio (relating width to height) varied on a trial-by-trial basis, were rated as looking wider when a /woo/ sound was presented at the same time, and as looking taller when a /wee/ sound was presented instead. By contrast, the perceived shape was not affected by other natural sounds such as birds or engine sounds, showing that a specific crossmodal effect was at stake between these sounds and these shapes. On the one hand, these findings add to a growing body of evidence demonstrating that audio-visual correspondences can have perceptual (as well as decisional) effects (see Parise & Spence, 2012; Deroy & Spence, 2013, for a review). On the other hand, the results concerning sound-shape correspondences add a new puzzle to the list of arbitrary-looking crossmodal matching tendencies whose origin is hard to explain.

The puzzle is all the more general in the case of shape that the shape-sound correspondences have a wide set of documented effects and applications. Besides the aforementioned bias in shape perception, they are shown to facilitate language learning (Imai et al., 2008) and to be exploited in various audio-visual mapping technologies such as music visualization software representing sounds as shapes or sensory substitution devices encoding shapes as sounds (see Deroy & Auvray, 2012).

In the present paper, we suggest that the study of these multimodal associations surrounding shape can be related to the more classical cases of crossmodal transfer of shape between vision and touch documented as part of Molyneux's question (part 2). We review the dominant explanations offered to explain shape-sound correspondences, in terms of conceptual mediation (part 3) and innate hyper-connectivity which is not eliminated by perceptual learning (part 4), before arguing that the hypotheses of associative learning and common neurological representations, proposed to explain the tactile-visual Molyneux's shape transfer can also explain the shape-sound crossmodal matchings. In conclusion, we stress that the hypotheses currently investigated for shape matchings in touch and vision benefit from being extended to the more arbitrary-looking cases of matchings shapes between audition and vision, thereby stressing the multimodal dimension of shape.

## 1. A new Molyneux problem

Arbitrary-looking crossmodal matchings, as they are called (Maurer & Mondloch, 2005; Spence & Deroy, 2012b), can be defined as tendencies to associate distinct sensory features that do not obviously co-occur in experience or in the environment. For instance, moving away from sound-shape pairings for a moment, the tendency to pair higher-pitched sounds with brighter visual surfaces is also shown to be present in adults (Marks, 1974) and in infants (Maurer et al., 2006). So is the tendency to match higher frequency sounds with higher visual locations (e.g. Evans & Treisman, 2010; Spence, 2011 for a review). These pairings occur although brighter objects and animals do not (at least straightforwardly) emit higher pitched sounds than their darker counterparts, and although higher pitched sounds do not regularly come from higher locations in space. The same lack of environmental grounding holds for the correspondence between shapes and sounds: Unless it should turn out that angular objects give rise to sounds that are relevantly different from rounded objects when, for example, they are explored haptically (e.g. Guzman-Martinez et al., 2012), there seems



to be no straightforward environmental correlation between shapes and sounds of objects either.

Crossmodal correspondences between sounds and shapes (or between pitch, brightness and elevation) are difficult to square with the currently popular view that crossmodal associations need to be learned out of the natural multisensory statistics of the environment (see Spence, 2011). Their origin therefore prompts a series of questions. How do such crossmodal correspondences come to be present in humans and other animals? Do they have any ecological value? Determining here whether these sound-shape associations are innate (Ludwig et al., 2011; Maurer & Mondloch, 2005; Maurer et al., 2012) or acquired; and in this case, determining how they are acquired (see Martino & Marks, 1999; Spence, 2011; Walker et al., in press) raise, as we shall see, a new Molyneux's problem, which teaches us new lessons on the multimodal aspect of shapes.

The core of Molyneux's problem, raised initially by Molyneux back in the 17th century, in the heat of the rationalist-empiricist controversies (Locke, 1690; see also Morgan, 1977) is still very much relevant today (e.g., Held et al., 2011, Streri, 2012). The question is to determine whether the crossmodal matching observed between felt and seen shapes at a very early age is acquired through exposure and associative learning, or whether it pre-exists exposure instead. To put it in a philosophical way, the question consists in deciding whether the crossmodal matching of shapes is *a priori* or *a posteriori*. To put it in a psychological way: is the tactile-visual connection for shapes innate / hardwired or acquired?

All past and current replies to Molyneux's problem have been framed on the basis that the matching between tactile and visual shapes targets one and the same environmental property (that is, shape is viewed as an objective or primary quality. Note that Berkeley (1948) is one of the rare philosophers who seems to have accepted that tactile shapes and visual shapes can constitute different objective properties). This objective grounding is what gives the crossmodal matching of tactile shapes and visual shapes a form of necessity and rationality of interest to philosophers.

Now, necessity, rationality and objectivity are what become problematic when we turn to arbitrary crossmodal matchings between sounds and shapes; as they obviously do not target one and the same environmental feature. Certain shapes do not necessarily go with certain sounds. For instance, associating the sound 'Bouba' to a rounded rather than to an angular shape looks irrational and this association does not seem to inform us about an objective regularity. So why would we pair sounds to shapes? Due to these key differences, the mainstream proposals developed for the Molyneux-type of crossmodal associations have not been thought to be relevant to address this question.

The very fact that the crossmodal correspondences between shapes and sounds is called arbitrary comes from the fact that scientists have had a hard time pinning down a regular environmental correlation between the property of being of a certain shape and the property of emitting a certain sound's pitch. Even harder to explain are crossmodal correspondences between shapes and flavours (Deroy & Valentin, 2011) or between symbolic shapes and smells (Seo et al., 2010) which also do not receive a straightforward explanation as internalised statistics of the environments. These other matchings might deserve a separate treatment, but they stress the crux of the problem: If shapes and the other properties are not necessarily or regularly correlated, how could these matchings be learned by association? And if they are not learned by exposure, how could one make sense of the fact that we have evolved to have hard-wired or *a*

*priori* connections between the representations of shapes and these apparently unrelated properties in our mind / brain?

The competing options that have been recently proposed to explain arbitrary crossmodal matchings between shapes and sounds, as we shall see below, recycle the ones that were once proposed for Molyneux's case but that were subsequently rejected. On the one hand, the idea, initially proposed for Molyneux's cases (see Locke, 1690; or Morgan, 1977 for a review) is that matchings across sensory modalities take place through an association made at the level of ideas or concepts and, on the other hand, the idea, that they are fully present at birth (i.e. that they are *a priori*, see Kant, 1998).

## **2. Sounds-shapes correspondences as conceptually mediated**

The idea that crossmodal matchings require a conceptual mediation is very much the way Molyneux's cases were discussed at the times of Locke and Berkeley, when the connection was supposed to be established between the 'idea' of shape prompted by vision and the 'idea' of shape prompted by touch. The idea of a conceptual mediation is however no longer considered appropriate in order to explain early crossmodal matchings of visual and tactile shapes. However, in the case of arbitrary crossmodal matchings, this hypothesis is pursued by a growing number of researchers: Correspondences between pitch, brightness, and angularity, for instance, have been explained by the cognitive capacity that observers have to represent various sensory features, or dimensions, on a common scale (Martino & Marks, 1999; Walker et al., 2012), to metaphorically map one conceptual domain onto another (Shen, 1997; Shen & Eisemann, 2008; Williams, 1976) or to reason analogically (Premack & Premack, 2003; see also Deroy & Spence, 2013; Spence, 2011, for a discussion). Now, the main problem for these conceptual solutions comes from explaining the presence of crossmodal matchings at a very early age (e.g., as early as 4 months, for shapes and sounds, see Orztuck et al., 2012) and the difference between the neurological activations noticed for crossmodal correspondences and semantic or analogical reasoning (Sadhigani et al., 2009).

## **3. Shape-sound correspondences as remnants of non-functional innate connections**

The nativist idea that crossmodal matchings could be present from birth has been eliminated – at least in the case of non-arbitrary matchings – for a long time in favour of the less radically nativist claim that they come from innate learning mechanisms guided by amodal or redundant representations of time, space, and intensity in the brain (see Bahrack & Lickiter, 2012, for a review). The strong nativist option is however still very much present when it comes to explaining arbitrary crossmodal correspondences as shown by the growing popularity of what is called the 'neonatal synaesthesia hypothesis' (see Maurer et al., 2012, for a review). The idea here is that these correspondences come from a lack of differentiation of the infant's perceptual apparatus, and persist into adulthood due to of a lack of pruning or inhibitory feedback

of some of these non-functional connections (Maurer & Mondloch, 1995; Maurer et al., 2012).

Now, there are good reasons not to go back to strong nativist hypotheses, even to explain the arbitrary crossmodal matchings evidenced in infants. The putative functional role of arbitrary crossmodal matchings as coupling priors in multisensory learning (Ernst, 2007; Spence, 2011) and multisensory integration (Parise & Spence, 2012), or as a kind of crossmodal Gestalt grouping principle (namely, a kind of crossmodal grouping by similarity; see Spence, submitted), together with neurological differences (Sadaghiani et al., 2009; Spence & Parise, in press), are sufficient to distinguish them from non-functional associations that can exist in synaesthetes (no matter whether they are adults or children; see Ward, 2012). This adds to the fact that nativist explanations in general are now hard to support in face of the demand that innate traits are traced back to their genetic encoding (a demand which is not easy to meet for most nativist hypotheses, see Lewkowicz, 2011).

#### **4. Updating the associative learning and common coding hypotheses to explain sound-shape correspondences.**

In this section, we want to argue that the alternative to explain arbitrary crossmodal correspondences either by late conceptual mediation or as being innate is wrongly limited. A first step here consists in stressing that explanations in terms of statistical learning and/or common neural coding have been too swiftly excluded.

The assumption that pairings – between, for instance, angularity and high-pitched sounds – are not regularly experienced by infants is more of an ungrounded assumption. It rather appears to be the default conclusion once one cannot come up with a plausible environmental source for the correlation. It should be more thoroughly investigated by taking into account precise measurements of exposure. Audiovisual correspondences between shapes and sounds might also come from a specific domain, namely speech. The mouth movements observed when someone utters speech sounds like ‘Takete’ or ‘wee’ are more stretched (angular / narrow) than the wider rounded movements observed when one utters ‘Maluma’ or ‘woo’; suggesting a regular correlation between pitch and shapes. This restores the plausibility of an associative learning account, especially compatible with the idea that infants are particularly attentive to face / voice or mouth / sounds in the first months of their life (see the perceptual narrowing hypothesis, Lewkowicz, 2002).

The second assumption that the neurological representation of visual brightness and auditory pitch cannot have anything in common also appears to rely on a predetermined view of what the legitimate common amodal representations in the brain are (i.e., space, time – plus or minus number / magnitude and quantity / intensity; see Marks, 1978). This assumption does not consider other possibilities which are getting investigated in recent work in cognitive neurosciences, that movement (Held et al., 2011) and embodiment could act as common sensibles (note that movement was considered as such by Aristotle and Locke).

Once related to speech, the correspondences between sound and shape can also be explained not merely in terms of audiovisual associations, but also in terms of audio-motor associations, linking the sounds that one hears to the automatic articulatory

movements generated when listening to speech (Galantucci, Fowler, & Turvey, 2006). If the latter account were to be correct, this crossmodal correspondence would then become embodied (Pezzulo et al., 2011), grounded in sensorimotor associations, rather than being based on an external association between two sensory experiences, whose resemblance would be processed in an amodal manner.

One way to distinguish between the statistical and embodied accounts here would be to test whether this correspondence exists only in cases or in species where the vocalising follows the take-te-sharp mouth movements rule. Note that this can be contrasted with the correspondence between the sound-size of the source which can be found across species, independently of their rules of vocalization (see Ludwig et al., 2011).

It will further be interesting to determine whether the sound-shape and sound-size crossmodal correspondences are related, and whether the latter has multiple origins (perhaps originating both in external and embodied underlying factors). Understanding the role of embodied vs. external associations would certainly help to link Sweeny et al.'s (2012) results to others showing that the shapes we see - and respond to - can also influence the pitch (or fundamental frequency) of the speech sounds we utter (Parise & Pavani, 2011) or that making a mouth movement (consistent with 'ba' or 'da') can give rise to a McGurk effect (McGurk & MacDonald, 1976) when listening to speech sounds, just as when actually viewing someone else's mouth movements uttering those sounds (see Sams, Mottonen, & Sihvonen, 2005).

## 5. Conclusion

In this concluding section, we want to insist on the importance of focusing on shape-sound correspondences when thinking about shapes, especially in a multi-disciplinary approach. From a global / philosophical perspective, these correspondences encourage a broadening of the investigation of Molyneux's problem, initially focused on tactile and visual shapes, to more contingent associations which can come to matter as much for linguistic and perceptual behaviour. Interestingly, the associative and commonality hypotheses framed here to account for correspondences between shapes and auditory attributes are also at the moment pursued for 'non-arbitrary' matchings of visual and tactile shapes, raising important questions as to how these two shapes might interact, and how situations of single vs. distinct properties can come to differ.

From a more specific and empirical perspective, crossmodal correspondences between shapes and sounds have a role in language acquisition and linguistic intuitions (Imai et al., 2008). They can also explain the use of crossmodal adjectives to talk about sounds (e.g., sharp sounds). But mostly, as we want to highlight, they show all their importance when thinking about the optimization of auditory-visual translations, be it the 'auditory' translation of visual shapes – as in sensory substitution through devices which aim at compensating the loss of sight through a coding / decoding device, such as the vOICE (Meijer, 1992) or the Vibe (Hanneton et al., 2010; see also Auvray & Myin, 2009, for a review) or the visual translation of sounds; for instance as in musical composition software.

## References

- Auvray, M., & Myin, E. (2009). Perception with compensatory devices. From sensory substitution to sensorimotor extension. *Cognitive Science*, 33, 1036-1058.
- Bahrack, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. J. Bremner, D. J. Lewkowicz & C. Spence (Eds.), *Multisensory development* (pp. 183-206). Oxford, UK: Oxford University Press.
- Berkeley, G. (1948–1957). *The Works of George Berkeley, Bishop of Cloyne*. A.A. Luce and T.E. Jessop (eds.). London: Thomas Nelson and Sons. 9 vols.
- Bien, N., ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size: Crossmodal binding in pitch-size synesthesia: A combined TMS, EEG, and psychophysics study. *NeuroImage*, 59, 663-672.
- Bremner, A., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K., & Spence, C. (2013). Bouba and Kiki in Namibia? Western shape-symbolism does not extend to taste in a remote population. *Cognition*, 126, 165–172.
- Deroy, O., & Auvray, M. (2012). Reading the world through sensory substitution devices. *Frontiers in Theoretical and Philosophical Psychology* (DOI: 10.3389/fpsyg.2012.00457).
- Deroy, O., Crisinel, A., & Spence, C. (accepted). Crossmodal correspondences between odours and contingent features: Odours, musical notes, and arbitrary shapes, *Psychonomic Bulletin & Review*.
- Deroy, O. & Spence, C. (2013). Why we are not all synaesthetes (not even weakly so), *Psychonomic Bulletin & Review*, <http://dx.doi.org/10.3758/s13423-013-0387-2>
- Deroy, O., & Valentin, D. (2011). Tasting shapes: Investigating the sensory basis of cross-modal correspondences. *Chemosensory Perception*, 4, 80-90.
- Ernst, M. O. (2007). Learning to integration arbitrary signals from vision and touch. *Journal of Vision*, 7, 1-14.
- Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10, 1-12.
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13, 361-377.
- Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. *Current Biology*, 22, 383-388.
- Hanneton, S., Auvray, M., & Durette, B. (2010). The Vibe: A versatile vision-to-audition sensory substitution device. *Applied Bionics and Biomechanics*, 7, 269-276.
- Held, R., Ostrovsky, Y., de Gelder, B., Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011). The newly sighted fail to match seen with felt. *Nature Neuroscience*, 14, 551-553.
- Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, 109, 54-65.
- Kant (1998). *Critique of Pure Reason*. Cambridge: Cambridge University Press.
- Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.
- Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York: Liveright Publication.
- Kovic, V., Plunkett, K., & Westermann, G. (2009). The shape of words in the brain. *Cognition*, 114, 19-28.

- Lewkowicz, D. J. (2002). Heterogeneity and heterochrony in the development of intersensory perception. *Cognitive Brain Research*, 14, 41-63.
- Lewkowicz, D. J. (2011). The biological implausibility of the nature-nurture dichotomy and what it means for the study of infancy. *Infancy*, 16, 331-367.
- Locke, J. (1690). *An essay concerning human understanding*. London.
- Ludwig, V. U., Adachi, I., & Matzuzawa, T. (2011). Visuo-auditory mappings between high luminance and high pitch are shared by chimpanzees (*Pan troglodytes*) and humans. *Proceedings of the National Academy of Sciences USA*, 108, 20661-20665.
- Meijer, P. B. L. (1992). An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 39, 112-121.
- Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch and loudness. *American Journal of Psychology*, 87, 173-188.
- Marks, L.E. (1978). *The unity of the senses: Interrelations among the modalities*. New York: Academic Press.
- Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, 28, 903-923.
- Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, 10, 61-65.
- Maurer, D., & Mondloch, C. J. (2005). Neonatal synaesthesia: A reevaluation. In L. C. Robertson & N. Sagiv (Eds.), *Synaesthesia: Perspectives from cognitive neuroscience* (pp. 193-213). Oxford: Oxford University Press.
- Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound-shape correspondences in toddlers and adults. *Developmental Science*, 9, 316-322.
- Maurer, D., Gibson, L.C., & Spector, F. (2012). Infant synaesthesia. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 239-250). Oxford, UK: Oxford University Press.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices, *Nature*, 264 (5588), 746-748.
- Morgan, M. (1977). *Molyneux's question*. Cambridge: Cambridge University Press.
- Ozturk, O., Krehm, M., & Vouloumanos, A. (2012). Sound symbolism in infancy: Evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*. DOI 10.1016/j.jecp.2012.05.004.
- Parise, C., & Pavani, F. (2011). Evidence of sound symbolism in simple vocalizations. *Experimental Brain Research*, 214, 373-380.
- Parise, C., & Spence, C. (in press). Audiovisual correspondences in the general population. In Simner, J. (ed.) *Oxford Handbook of Synaesthesia*. Oxford: Oxford University Press.
- Peiffer-Smadja, N. (2010). Exploring the bouba/kiki effect: A behavioral and fMRI study. Unpublished Ms Thesis, Université Paris V – Descartes, France.
- Pezzulo, G., Barsalou, L.W., Cangelosi, A., Fischer, M.H., McRae, K., & Spivey, M.J. (2011). The mechanics of embodiment: A dialog on embodiment and computational modelling. *Frontiers in Psychology*, 2, 5.
- Premack, D., & Premack, A. J. (2003). *Original intelligence: Unlocking the mystery of who we are*. New York: McGraw-Hill.
- Ramachandran, V. S., & Hubbard, E. M. (2001a). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies*, 8, 3-34.

- Ramachandran, V. S., & Hubbard, E. M. (2001b). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society London B*, 268, 979-983.
- Sadaghiani, S., Maier, J. X., & Noppeney, U. (2009). Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *Journal of Neuroscience*, 29, 6490-6499.
- Sams, M., Mottonen, R., & Sihvonen, T. (2005). Seeing and hearing others and oneself talk. *Cognitive Brain Research*, 23, 429-435.
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12, 225-239.
- Seo, H.-S., Arshamian, A., Schemmer, K., Scheer, I., Sander, T., Ritter, G., & Hummel, T. (2010). Cross-modal integration between odors and abstract symbols. *Neuroscience Letters*, 478, 175-178.
- Shen, Y. (1997). Cognitive constraints on poetic figures. *Cognitive Linguistics*, 8, 33-71.
- Shen, Y., & Eisenamn, R. (2008). Heard melodies are sweet, but those unheard are sweeter: Synaesthesia and cognition. *Language and Literature*, 17, 101-121.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971-995.
- Spence, C., & Deroy, O. (2012a). Hearing mouth shapes: Sound symbolism and the reverse McGurk effect. *i-Perception*, 3, 550-552.
- Spence, C. & Deroy, O. (2012b) Crossmodal correspondences: Innate or learned? *i-Perception*, 3, 316-318.
- Spence, C., & Parise, C. V. (2012). The cognitive neuroscience of crossmodal correspondences. *i-Perception*, 3, 410-412.
- Streri, A. (2012). Crossmodal interactions in the human newborn: New answers to Molyneux's question. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 88-112). Oxford, UK: Oxford University Press.
- Sweeny, T. D., Guzman-Martinez, E., Ortega, L., Grabowecky, M., & Suzuki, S. (2012). Sounds exaggerate visual shape. *Cognition*, 124, 194-200.
- Walker, L., Walker, P., Francis, B. (in press). A common scheme for cross-sensory correspondences. *Perception*.
- Ward, J. (2012). Synaesthesia. In B. E. Stein (Ed.), *The new handbook of multisensory processes* (pp. 319-333). Cambridge, MA: MIT Press.
- Williams, J.M. (1976). Synesthetic adjectives: A possible law of semantic change. *Language*, 52, 461-478.

# Structure, Similarity and Spaces

Sandro Rama FIORINI<sup>a,1</sup>, Mara ABEL<sup>a</sup> and Peter GÄRDENFORS<sup>b</sup>

<sup>a</sup> *Institute of Informatics, UFRGS, Brazil*

<sup>b</sup> *Lund University Cognitive Science, Lund University, Sweden*

**Abstract.** Much of the discussion about shape representation during the last two decades was fundamentally related to questions about the representation of parts. Inspired by the cognitive processes governing how people represent and think about parts, we give a brief summary of our framework for representing part structures. In particular, we are interested in the role of similarity and prototype effects in this context.

**Keywords.** Part-whole relations, similarity, conceptual spaces

## Introduction

Humans seem to be prone to divide the complex shape of objects into parts. In seeing a cat, we divide its overall shape in some more-or-less well defined parts, such as the head, trunk, tail and legs. We can then use this information to recognize and think about cats. It seems that structure is intrinsically related to our everyday notion of shape. That leads us to a broader question: what are the cognitive phenomena that allow us to represent and think about parts of object as a whole, and not just parts of their shapes? In this paper, we introduce a novel way to represent the relation between parts and wholes that takes into consideration some of these phenomena.

In an influential work, Farah [4] carried out a meta-study about different kinds of agnosia in humans and proposed that two processes participate in object recognition. Object recognition can be *structural*, where recognition is achieved by the identification of parts of the object and its internal structure. On the other side, recognition can be *holistic*, based on the global characteristics of the object, such as overall shape. According to Farah, both processes work together in recognitions of broad categories of entities (e.g. faces, objects and written words). It is important to note that, as pointed by Peissig and Tarr [7], the structural *versus* holism problem is correlated, but independent of the recurrent question whether object (shape) recognition is model- or view-based (i.e. [2,3]).

If we assume that holistic and structural processes are necessary for object recognition, therefore it is reasonable to expect that both require an underlying conceptual structure conveying holistic and structural information. When it

---

<sup>1</sup>Corresponding Author: Sandro Fiorini, Institute of Informatics, UFRGS, P.O. Box 15064, CEP 91501-970, Porto Alegre-RS, Brazil; E-mail: srfiorini@inf.ufrgs.br. We thank CAPES and Petrobras project PB PRH-17 for the support in this work and Joel Carbonera for the comments.



comes to conceptual structures for object recognition, similarity effects seem to play a fundamental role (c.f. [3,1]). The general notion is that recognition could be reduced to judgements of similarity between perceptual input and internal representations. The question we are trying to answer is how a holistic/structural representation framework that supports similarity judgements should look like. We base our efforts on the Theory of Conceptual Spaces [5], a representation framework that embeds the notions of concept similarity and prototypes. Our general approach is to extend conceptual spaces so that it becomes more suitable for the representing holistic and, specially, structural information about concepts and objects. It should provide the grounds for novel computational approaches to concept representation based on holistic and structural similarity.

## 1. Conceptual Spaces

Gärdenfors' Theory of Conceptual Spaces [5] puts forward a new way for representing concepts using geometrical and topological structures, which complements symbolic and connexionist approaches. Given the available space, we present just a brief introduction, but it is enough to say that the Theory is based on the notions of concept similarity and prototypes. Put it simple, a conceptual space is a space in the mathematical sense, where objects are points and concepts are regions or sets of regions. If this space has a well-defined metric, then it is possible to tell the similarity between objects (and concepts) by measuring the distance between them: further objects are apart, less similar they are. The dimensions of a conceptual space have a special meaning: they denote the features — or qualities — through which entities can be compared and are frequently grounded in perception. Good examples are hue, mass, height, etc. Certain quality dimensions always co-occur, forming subspaces called quality domains. Examples of quality domains are colour, shape, taste, etc. Complex concepts are defined as set of regions in many quality domains. For instance, the concept of apple can be defined as a combination of the regions green and red in the colour domain, plus the cycloid regions in the shape domain, plus the sweet and acid regions in the taste domain and so on. An individual apple is represented by a single point (or vector) in the multi-dimensional space formed by all quality domains of apple and that is close (similar) enough to the regions that form the concept of apple. A type of apple is formed by subregions of apple.

## 2. Structure Spaces

We are interested in using conceptual spaces to represent holistic and structural information about concepts and objects. The holistic portion of an entity can be seen as its usual features: colour, shape, weight, etc. These can be readily represented as regions in quality domains. However, representing structural information is far from trivial. In doing so, we are fundamentally interested in describing the partonomic relations between parts and wholes. The Theory of Conceptual Spaces does not provide a complete solution for representing relations in general;

it simply suggests that they could be represented by a (Cartesian) product between the relata. We take this basic notion and develop it further to account for structural information.

Structural information can be represented in conceptual spaces through what we call *structure spaces*. A concept are represented as regions in a conceptual space; a structure space is a subspace of it, where part structures are described. A vector in a structure space denotes a particular configuration of parts of an individual. That is, a single vector encode the information about what parts compose a whole and also about how these parts are related to whole. Similar configurations of parts are close together in this space. For a single vector to convey all this information, much of it is naturally allocated to the dimensions. Given a concept  $C$  and a set of concepts  $P_1, \dots, P_n$  denoting parts of  $C$ , then we can generally define the structure space containing  $C$  as the product space of the quality domains of  $P_1, \dots, P_n$  and  $n$  quality domains denoting specific structure information about each part  $P_i$ . We call these domains *structure domains*; they represent information such as the displacement of the part in the whole, part quantification and so on. The actual structure space of  $C$  is the product space of the regions that define  $P_1, \dots, P_n$ , plus regions in the structure domains. For instance, the structure space of Apple could be formed by the product of Core, Flesh, Seed and Stein, plus regions denoting the general positioning each part in an object-centred coordinate system. A vector in the structure space of  $C$  denotes a particular apple-structure: a combination of individual parts, each with a specific value for colour, shape, taste and so on. Close points in this space represent similar apple-structures. The combination of regions of each part in the product restricts what are the valid individual components of an apple. More importantly, the structure space can be further divided into specific regions defining types of apple-structures; e.g., the concept of an apple with acid flesh and short stein.

This basic formulation of our framework might raise questions, such as problem of co-determination between holistic and part qualities; the role of parts in taxonomies (c.f. Tversky [8]), the representation of the many kinds of partonomic relations; the question about dependent and essential parts; computational feasibility and so on. To all of those we have at least partial answers, but due to the limited space, we shall touch upon the issue we consider most critical: the problem of transitivity.

Wholes have parts, which can also have parts and so on. This might become an issue when we define a whole as a product of parts: given parts also can have parts, complex wholes could soon become multidimensional monsters; the structure space of the concept Universe would be impossible to describe fully. In more formal terms, we could say that the structure space of a whole might become the transitive closure of its parts. In order to solve this problem, we first assume that the part relation is not essentially transitive. This position contrast with formal theories of parts, such as Classical Mereology, but it has becoming increasingly common in recent years (c.f. [6]). Instead, we take a more cognitive stance; experience and perception are the sole determinants of which parts *directly* compose a whole. For instance, what determines that the person's heart is not part of the company where she works in is the fact that there is no use for such conceptualization in the actual context. However, if there is a change in

context (e.g. parts of employee body becoming property of the company), we can easily adapt our conceptualization. We do not assume any hard *a priori* ontological distinction on parts and wholes, for it might hinder the plasticity of the representation. Instead, we provide a way in which plasticity could operate in our framework, by improving the definition of structure spaces. We introduce the notion of *dimensional filter*. A dimensional filter is a conceptual operator that projects a subset of the quality domains of a concept onto a smaller subset; it “selects” relevant domains of another concept. We can then redefine a concept as a product space of filtered parts, where just the relevant domains are selected to compose the structure space of the whole. The filtering (i.e. projection) is controlled by processes like attention and context. For instance, a combustion engine can be part of a car or part of an electricity generator. The quality domains of the engine that are relevant for car are related to its characteristics as a car mover. So, the projection of combustion engine into the structure spaces of car carries only some of its more relevant domains. This scheme solves the dimensional explosion by providing a way in which the transitive closure can be avoided; parts of parts that are not relevant for the whole can be filtered out.

### 3. Final Remarks

We are now developing a mathematical formulation of structure spaces based on metric spaces. This should pave the way for computational applications. We are also investigating the use of structure spaces in robotics and geology. Some auto-localization algorithms for robots employ similarity reasoning to compare its surroundings with its internal map. This comparison could benefit from a representation that allows structural similarity matching. In petroleum geology, an analogous problem of structural similarity exists in the task of matching geological structures in different exploration wells, to which no satisfactory computational solution yet exists. In the same way, structure spaces could help solve this problem by providing a principled way in which sequences of features can be compared.

### References

- [1] R. G. Alexander and G. J. Zelinsky. Effects of part-based similarity on visual search: the frankenbear experiment. *Vision research*, 54:20–30, Feb. 2012. PMID: 22227607.
- [2] I. Biederman. Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94:115–117, 1987.
- [3] S. Edelman. Representation is representation of similarities. *Behavioral and Brain Sciences*, 21(04):449–467, 1998.
- [4] M. J. Farah. Is an object an object an object? Cognitive and neuropsychological investigations of domain specificity in visual object recognition. *Current Directions in Psychological Science*, 1(5):164–169, 1992.
- [5] P. Gärdenfors. *Conceptual Spaces: The Geometry of Thought*. The MIT Press, Cambridge, Massachusetts, 2000.
- [6] I. Johansson. Formal mereology and ordinary language - reply to varzi. *Appl. Ontol.*, 1(2):157–161, Apr. 2006.
- [7] J. J. Peissig and M. J. Tarr. Visual object recognition: Do we know more now than we did 20 years ago? *Annual Review of Psychology*, 58:75–96, Jan. 2007.
- [8] B. Tversky. Parts, partonomies, and taxonomies. *Dev. Psychol.*, 25(6):983–95, 1989.

# Local Qualities, Quality Fields, and Quality Patterns: A Preliminary Investigation

Nicola GUARINO

*ISTC-CNR Laboratory for Applied Ontology, Trento*

## Abstract

**Abstract.** When we describe the shape of certain entities, like a vase or a river, we refer to their qualities in different ways. A river has usually a definite length, but its width varies with the distance from the source, typically getting higher towards the end. Similarly, a vase has a definite height, but its width may vary, reflecting a certain pattern that often marks a particular style. So, at least for certain entities, quality kinds such as length, height and width don't behave in the same way: length or height just inhere to these objects with no need of further qualification, while width requires a spatial localization in order to be determined.

In this paper I would like to explore the way qualities of things behave with respect to the parts of such things. Building on the notion of individual quality introduced in the DOLCE ontology, I will introduce the new notions of local quality, quality field and quality pattern, stressing their cognitive role in many practical situations. I will argue that an expression like "the river's width" or "the depth of the sea" actually refers to a quality field, and not to an individual quality. Quality fields will be used to introduce the further notion of *quality pattern*, and to analyze the distinction between *variation* and *change*.

**Keywords.** Ontology, quality, local quality, global quality, property, field, pattern

## 1. Introduction

When we describe the shape of certain entities, like a vase or a river, we refer to their qualities in different ways. A river has (more or less) a definite length, but its width varies with the distance from the source, typically getting higher towards the end. Similarly, a vase has a definite height, but its width may vary, reflecting a certain pattern that often marks a particular style. So, at least for certain entities, quality kinds such as length, height and width don't behave in the same way: length or height just inhere to these objects with no need of further qualification, while width requires a spatial localization in order to be determined.

We shall say that length and height, in these examples, behave as *global qualities*, while width behaves as a *local quality*. A local quality of a certain object is a quality which actually inheres to a *part* of that object, but, despite this fact, is somehow considered, from the cognitive point of view, as a quality of the whole object: so, we rarely say "the width of this river stretch is 100 meters", but we prefer to say "the river's width is 100 meters *here*". Analogously, we say "the depth of the Adriatic Sea is much higher along the Croatian coast than along the Italian coast", referring to "the river's width" or "the sea's depth" as one single entity, although, so to speak, spread out in space. Indeed, in many simple cases, we describe the *qualitative shape* of a certain object in terms of the behavior of a local spatial quality along a certain dimension.

Of course, the distinction between global and local qualities is very general, and goes much beyond purely spatial qualities: consider for instance the mass or volume of a physical object vs. its density or its temperature, or the duration of a rain vs. its intensity. In all these cases, we observe different ways qualities of things behave with respect to the parts of such things. The main purpose of this paper is to explore this phenomenon, which I will call *mereological behavior of qualities* (or, more in general, of *properties*, as we shall see), providing a formal account of local qualities and analyzing its practical implications.

## 2. Inclusive and exclusive properties

Looking at the philosophical literature, the phenomenon we have described appears to be connected to a more general one, concerning the mereological behavior of properties. A classic distinction in this respect is that between homoeomerous and anomoeomerous properties, based on whether or not a property holding for a whole also holds for all its parts, and discussed in particular by Armstrong [1]. Ingvar Johansson [2] builds on this work in the light of the Johnson's distinction between determinates and determinables [3], focusing his attention to the case of determinate properties belonging to the same determinable, and to the ontological nature of patterns like for instance a distribution of colored areas on a surface [4]. This is, at least in my knowledge, one of the few works addressing in some detail the mereological behavior of what I call qualities<sup>1</sup> (i.e. colors, lengths, temperatures), and not just that of generic properties, so I think it is a good starting point in our analysis<sup>2</sup>. Johansson proposes the following distinction among determinates:

- (1) A determinate property is *inclusive* if and only if each possible part of its instances instantiates some other property under the same determinable.
- (2) A determinate property is *exclusive* if and only if each possible part of its instances instantiates this very same property (under the same determinable).

As an example of inclusive property, Johansson brings a volume-determinate: if an object has a volume of 100 cm<sup>3</sup>, then all its (proper) parts must have a different volume. So every instance of an inclusive determinate necessarily includes other (different) determinates for its parts<sup>3</sup>. On the contrary, he notes that a color-determinate is exclusive: if an object is (*homogeneously*) red, different determinates (under the same determinable) are excluded.

In my view, the homogeneity proviso in the latter example is illuminating, since it gives evidence of a peculiarity of some determinables (like color, density, or temperature): they are prone to have their determinates arbitrarily distributed in the region they are defined. They have therefore a *local behavior*. This makes it difficult to understand the meaning of simple statements like *this car is red*, or *the room's temperature is 20 °C*, so that, in many cases, the ascription of such determinates to an extended object is the result of an implicit convention (the car color is the body's color) or an average operation, and very rarely results from an homogeneous distribution.

To account for possible non-homogeneous distributions of determinates under the same determinable, Johansson introduces a variant to (2):

- (3) A determinate property is *semi-exclusive* if and only if some possible part of its instances instantiates this very same property (under the same determinable).

So, if we admit that the red determinate is semi-exclusive, then an object can be globally red (as a result of some convention) if it just has one red part while being locally yellow somewhere else, or it may be also the case that it has multiple local colors with no definite global color. We can conjecture therefore that semi-exclusivity is a formal property which is associated with local behavior.

---

<sup>1</sup> Note however that Johansson's notion of quality is different from mine, as will be clear in the following, since he adopts a realist ontology according to which qualities are conceived as universals, and include both substances and properties.

<sup>2</sup> To complete the picture, we should also consider the distinction between *extensive* and *intensive* properties, which has many variants in the literature. I will discuss it in the next section, since I believe that Johansson's work helps to understand it.

<sup>3</sup> For the sake of simplicity, I will assume there are no atoms in the domain, so every object has some proper parts.

Consider now the width of a river. Like for the color case, we can have a river stretch whose width is homogeneous, or another whose width is not definite, since it varies along the river flow. So width-determinates are semi-exclusive. However, differently from colors, we can't say that widths can distribute freely for all the parts of the river stretch, since the longitudinal parts are constrained to have a width lower than that of the river stretch. Indeed, besides being semi-exclusive, according to Johansson widths are also *semi-inclusive*:

- (4) A determinate property is *semi-inclusive* if and only if some possible part of its instances instantiates some other property under the same determinable.

This means that, in practice, there is a *direction* (across the river flow) according to which each part contributes to the width of the whole, but there is another direction (along the flow) according to which there is no such contribution. As we shall see, this observation will be useful to define the notion of local quality. For the time being, however, we must say that semi-exclusivity, alone, is not enough to fully capture the intuition behind the notion of local behavior, mainly because, according to the definitions above, semi-inclusivity and semi-exclusivity are not disjoint.

### 3. Extensive and intensive properties

In physics and chemistry, an extensive property (like having a certain volume) is such that, if it holds for a whole, every single part necessarily contributes to this fact, while an intensive property (like having a certain temperature or color) is not necessarily affected by the parts of its instances. In his book, Johansson uses the terms 'extensive' and 'intensive' in a different sense, based on Kant's distinction between extensive and intensive *magnitudes*. He observes that for Kant 'extensive' is synonymous of 'inclusive', but 'intensive' means just *quantifiable*, so that it is not opposite to 'extensive'<sup>4</sup>. I don't think this understanding of 'intensive' is useful for our purposes, so I will stick to the previous (still informal) definition.

I will rather suggest the following definition:

- (5) A determinate property is *extensive* if and only if it is necessarily semi-inclusive. A property is *intensive* otherwise.

Under this definition, an intensive property is just a property which does not impose any inclusivity constraints on its mereological behavior. All exclusive properties turn out to be intensive, as well as all semi-exclusive properties *which are not also semi-inclusive*<sup>5</sup>. An important aspect of this definition is its modal nature. A property is extensive or intensive depending on whether or not, because of its very nature, it forces the proper parts of its instances to be different. To me, this definition of intensivity finally captures our intuitions concerning the local behavior of a determinable: only intensive determinables (i.e., all of whose determinates are intensive) can admit arbitrary distributions within the parts of a given object.

As a final note, someone may observe that (5) is neutral with respect to an aspect usually considered as characteristic of extensive properties, namely additivity. Additivity however only concerns quantifiable properties, and I agree with Johansson that the basic distinction we are aiming at should be more general. In particular,

<sup>4</sup> In other words, on one hand 'intensive' is understood as opposite to 'extensive', while on the other hand it is understood as 'capable of having an intension'.

<sup>5</sup> In addition, a property like *mean temperature of 20° C*, which is neither semi-exclusive nor semi-inclusive (since it can hold both when all the proper parts of its instances have the same temperature and when no proper part has the same temperature), turns out to be intensive.

Johansson makes a very interesting example of a semi-inclusive determinable, namely *shape*, which is not additive but still would count as extensive under our definition.

#### 4. Local qualities

Let us now shift from determinates and determinables to individual qualities and quality kinds. Independently from any deep metaphysical commitments, I think that this move is pragmatically useful not only to provide an analysis consistent with the DOLCE ontology [5], [6], but also to better clarify our intuitions concerning local qualities. In DOLCE we distinguish between *individual qualities*, which inhere to specific individuals, and *qualia*, which are abstract entities representing what exactly resembling individual qualities have in common. Qualia resulting from comparable individual qualities are organized in quality spaces. We refer to individual qualities with expressions such as *the color of my car*, while we refer to qualia with simple terms like *red*. So the color of my car is different from the color of your car, even if they have exactly the same shade of red, i.e., they have the same quale<sup>6</sup>. If I paint my car in blue, the term *the color of my car* still denotes the same quality, but its quale is now different. So, in general, the relationship between qualities and qualia is temporalized.

From now on, I shall simply use the term *quality* to refer to an individual quality. In DOLCE, qualities are organized in *quality kinds*, i.e. *maximal classes of comparable qualities*. *Color* and *volume* are examples of quality kinds. Their instances are individual colors or individual volumes. The objects to which such qualities inhere are instances, respectively, of the *has-color* and *has-volume* determinables. So there is a correspondence between determinables and quality kinds, but quality kinds are not determinables. Similarly, there is a correspondence between determinates and qualia, in the sense that the objects whose qualities have the same quale inhere are instances of the same determinate.

In terms of individual qualities and quality kinds, (1) can be reformulated as follows<sup>7</sup>:

- (6) A quality  $q$  of kind  $Q$  with quale  $ql$  inhering to an object  $x$  is *inclusive* iff every proper part  $y$  of  $x$  has a quality  $q'$  of kind  $Q$  with quale  $ql'$  different from  $ql$ .

The definitions (2)-(4) above can be reformulated analogously, and will be omitted here for the sake of brevity. (5) however makes little sense at the level of individual qualities, and much more sense at the level of quality kinds, or, more in general, to classes of qualities belonging to the same kind:

- (7) A quality kind (or a class of qualities belonging to the same kind) is *extensive* iff, necessarily, all its instances (i.e., its individual qualities) are semi-inclusive. It is *intensive* otherwise.

Let us now go back to our initial question: *what is a local quality*? A first conjecture we can make is that *a local quality is an instance of an intensive quality kind*. So, looking back at our original example, the color of each part of the vase, being an instance of an intensive quality kind (color), appears to be a local quality.

<sup>6</sup> I must say I have no clear understanding about the best way to name the relationship existing between individual qualities and their qualia. In this paper I will simply use "has". Probably a more informative name would be "manifests as".

<sup>7</sup> For the sake of conciseness, I will abstract from time in these definitions, since time is not relevant to understand their rationale. Extending them to the temporal version is however a simple matter, taking into account that in DOLCE the part-of and the has-quale relation are temporalized when they concern endurants.

However, things are not so simple, since we want also to account for the local behavior of a river's width, and *width* is not an intensive quality kind, since it is semi-inclusive. Moreover, a fundamental question arises: *local to what?* Given a particular object, how can we tell what are *its* local qualities? For instance, think of our river example. Arguably, the river has many local widths, each inhering to a particular river stretch. But also each possible longitudinal part of the river has a width. Is that width a local quality *of the river*, or just of a strange part of it? From the cognitive point of view, I say it is not a local width of the river. Think of another example, perhaps more vivid: a vase with a handle. The vase has many local widths (at the top, at the bottom, in the middle...), as well as the handle has many local widths, but none of the local widths of the handle is a local width of the vase, despite the fact that the handle is a part of the vase.

To address this issue, let us first clarify that each local quality of an object inheres to some proper part of the object, *but not to the object itself*. So, when we say that a vase *has* many local widths, this is because of an *indirect* inherence relation, similar to the one holding in DOLCE between an event and the spatial location of its participants. The point now is that this indirect relation existing between the quality of a part and the whole doesn't hold for all parts: as we have seen, not every width of a part of the vase is a local width of the vase. I think there is a simple cognitive mechanism that explains this situation. When we consider a (non-inclusive) quality kind such as *width* as applied to a river, we implicitly introduce a set of *canonical parts* (namely those cut across the river flow, corresponding to river stretches) whose width counts as a *local width of the river*. Because of the way canonical parts are constructed, the class of all local widths of a river turns out to be intensive, despite the quality kind 'width' is not.

Similarly, when we apply 'width' to a vase, we consider as canonical parts those concerning the internal cavity, excluding the handle. Finally, when we apply 'color' to a car, we only consider the external parts of the body as canonical parts, while when we apply it to the vase we may consider as canonical all the (external) parts of the vase, including those of the handle.

In sum, we can conclude that local qualities are instances of *quality classes*, constructed ad hoc with the cognitive mechanism described above as a specialization of non-inclusive quality kinds. In general, such classes are not *rigid*, using OntoClean's terminology: the color of a part of the vase will keep its identity when the part is removed, although it will not be a local color of the vase any more. Local qualities don't form a new ontological category. Locality, for qualities, is just a special way to describe a whole (especially its *shape*) with reference to the qualities of its parts.

## 5. Quality fields and quality patterns

Consider now all local qualities of a certain kind that a certain object has, at a given time. For instance, think of all the local depths of the Adriatic Sea. I say that the mereological sum of these local qualities is a *quality field*. While a single local depth does not inhere to the whole Adriatic Sea, the whole depth field does, it inheres to it. Quality fields form therefore a new ontological category in the vast class of dependent entities. We are able to define them thanks to the introduction of the notion of local quality.

Once we introduce quality fields, we can give an exact denotation to expressions like "the river's width" or "the depth of the sea" which don't usually refer to an individual quality, unless we conventionally pick up a specific individual quality (say the maximal depth). I will say that these expressions denote a quality field. Indeed, on the basis of our previous analysis, we can safely assume that an expression of the form



“the Q of X” denotes a quality field if Q is not an extensive quality kind, unless a special convention is introduced for certain kinds of X.

Let us bring time into account, using the notion of quality field to analyse the distinction between *variation* and *change*. Synchronically, at a specific time, we can observe *variations* in the Adriatic Sea’s depth field when shifting our focus of attention from one part (e.g., the depth of the Italian side) to an other (e.g., the depth of the Croatian side). The whole quality field can also genuinely *change* in time, keeping its identity, when some of its individual qualities change their qualia. This is what happened in the last 2000 years, since the depth of the Italian coast is much lower than it used to be in the Roman age.

Consider now the Adriatic Sea’s depth field at the Roman age. It exhibits an individual spatial *pattern*, which is different from the pattern we can observe today. Such pattern reflects *the specific qualia distribution* of the individual qualities forming the depth field at that time. I define a individual pattern as an emerging entity *constituted* by (a part of) a quality field. It differs from a quality field since the actual qualia distribution is an essential property of the pattern, and is not essential to the quality field (whose qualia distribution can change in time). So individual patterns are frozen, they don’t change, although they can exhibit variations. Note that I am insisting using the term *individual* pattern to make it clear they are not abstract entities: two things may have exactly resembling, although distinct, individual patterns. In this case, I say that they will have the same *shape*, but this is another story.

Let us finally shift the attention to perdurants. They can have both global and local qualities, inhering to each temporal part. The duration of a rain is a global quality, while its intensity is a quality field. When we say “the rain intensity is high now” we are referring to the whole intensity field, whose value happens to be low in the present time interval, exactly like the width of the river happens to be high in a certain place. So, the introduction of local qualities and fields can help understanding apparent “changes” in events and processes, like those described in [7], without the need of introducing a new entity which is the subject of such change: a speed variation during a run or an increase in the river flow will be simply considered as a variation of a local quality (the run speed or the water flow) along the temporal dimension. This move may also eliminate (at least for these examples) the need to introduce so-called *relators* [8] in addition to events exemplifying a binary relation like *marriage*: a change in the peacefulness of the marriage will be simply described as a variation of a local quality of the marriage itself, with no need to introduce a separate relating entity.

In conclusion, the introduction of local qualities, fields and patterns seems to be a simple and powerful extension to the Dolce’s notion of quality, which allows us to formally account for the way we deal with spatial and temporal quality distributions. This is still not enough to fully account for the different *shapes* such distributions may exhibit, but I believe it contributes to understand what shapes are.

## References

- [1] D. M. Armstrong, *Universals and Scientific Realism, Volume 2: A Theory of Universals*. Cambridge University Press, 1978, pp. 1–169.

- [2] I. Johansson, *Ontological Investigations. An Inquiry into the Categories of Nature, Man and Society*. Routledge, 1989.
- [3] W. E. Johnson, "Logic - Part I," *Cambridge University Press*, pp. 1–304, 1921.
- [4] I. Johansson, "Pattern as an ontological category," N. Guarino, Ed. *Formal Ontology in Information Systems* Amsterdam, Oxford, Tokyo, Washington, DC: IOS Press (Frontiers in Artificial Intelligence and Applications), 1998, pp. 86–94.
- [5] S. Borgo and C. Masolo, "Foundational choices in DOLCE," in *Handbook on ontologies*, S. Staab and R. Studer, Eds. Springer, 2009, pp. 361–381.
- [6] C. Masolo, "The DOLCE Ontology," WonderWeb Deliverable D18, Dec. 2003. Available at <http://www.loa.istc.cnr.it/Papers/D18.pdf>.
- [7] A. Galton and R. Mizoguchi, "The water falls but the waterfall does not fall: New perspectives on objects, processes and events," *Applied Ontology*, vol. 4, no. 2, pp. 71–107, 2009.
- [8] G. Guizzardi, *Ontological foundations for structural conceptual models*. University of Twente, CTIT PhD Thesis Series, no. 05-74, 2005.



# Shape Perception in Chemistry

Janna HASTINGS <sup>a,b,c,1</sup>, Colin BATCHELOR <sup>d</sup> and Mitsuhiro OKADA <sup>e</sup>

<sup>a</sup>*Cheminformatics and Metabolism, European Bioinformatics Institute, Hinxton, UK*

<sup>b</sup>*Swiss Centre for Affective Sciences, University of Geneva, Switzerland*

<sup>c</sup>*Evolutionary Bioinformatics, Swiss Institute of Bioinformatics, Switzerland*

<sup>d</sup>*Royal Society of Chemistry, Cambridge, UK*

<sup>e</sup>*Department of Philosophy, Keio University, Tokyo, Japan*

**Abstract.** Organic chemists make extensive use of a diagrammatic language for designing, exchanging and analysing the features of chemicals. In this language, chemicals are represented on a flat (2D) plane following standard stylistic conventions. In the search for novel drugs and therapeutic agents, vast quantities of chemical data are generated and subjected to virtual screening procedures that harness algorithmic features and complex statistical models. However, in silico approaches do not yet compare to the abilities of experienced chemists in detecting more subtle features relevant for evaluating how likely a molecule is to be suitable to a given purpose. Our hypothesis is that one reason for this discrepancy is that human perceptual capabilities, particularly that of ‘gestalt’ shape perception, make additional information available to our reasoning processes that are not available to in silico processes. This contribution investigates this hypothesis.

Algorithmic and logic-based approaches to representation and automated reasoning with chemical structures are able to efficiently compute certain features, such as detecting presence of specific functional groups. To investigate the specific differences between human and machine capabilities, we focus here on those tasks and chemicals for which humans reliably outperform computers: the detection of the overall shape and parts with specific diagrammatic features, in molecules that are large and composed of relatively homogeneous part types with many cycles. We conduct a study in which we vary the diagrammatic representation from the canonical diagrammatic standard of the chemicals, and evaluate speed of human determination of chemical class. We find that human performance varies with the quality of the pictorial representation, rather than the size of the molecule. This can be contrasted with the fact that machine performance varies with the size of the molecule, and is of course impervious to the quality of diagrammatic representation.

This result has implications for the design of hybrid algorithms that take features of the overall diagrammatic aspects of the molecule as input into the feature detection and automated reasoning over chemical structure. It also has the potential to inform the design of interactive systems at the interface between human experts and machines.

**Keywords.** ontology, shape perception, cognition, spatial reasoning, logical reasoning, molecular graph

---

<sup>1</sup>Corresponding Author, e-mail: [hastings@ebi.ac.uk](mailto:hastings@ebi.ac.uk)

## Introduction

“A mind that has the ability to choose how it will represent a particular problem it needs to solve, choosing from a repertoire of representational capacities that include more analogical and more symbolic notations is more flexible, hence more ‘intelligent’ ” [32]

Organic chemists make extensive use of chemical diagrams for designing, exchanging and analysing the features of chemicals. In this language, chemicals are represented on a flat (2D) plane following standard stylistic conventions [12]. The use of diagrammatic languages to concisely convey information for humans to process is an essential component of many sciences. In biology, pathway diagrams convey information about biological processes [14]. A good visualization of scientific information facilitates rapid understanding and can thereby lead to novel insights not otherwise possible [19]. One such example is Category Theory in mathematics, in which the use of diagrams is essential for representing mathematical properties and proofs [30].

In the search for novel drugs and therapeutic agents, large quantities of chemical data are generated. Interacting with these data and sifting a relevant subset (for a given problem) from the sizeable background is an ongoing challenge. Tools such as the *molecule cloud* can give an overview of a chemical dataset by showing common scaffolds sized for how often they appear in the dataset [9]. Many features of chemical entities have relevance on whether a given molecule is suited to a given purpose. Algorithmic and logic-based approaches are able to efficiently compute certain of these features, such as the presence of specific atoms or functional groups, overall mass and charge [18]<sup>2</sup>. Algorithmic approaches can also gauge the overall shape of a molecule (at least in terms of delineating the outline of the three-dimensional space it fills) and calculate the mathematical similarity of that shape to that of other molecules or the reciprocity to potential binding sites [40,2,20]. Yet, many problems in chemical informatics remain difficult to efficiently automate over large molecular collections (e.g. finding maximal shared components between a set of molecular graphs [31], detecting all the cycles in a given molecule [3]).

In what follows, we focus on a class of problems that are known to be challenging for algorithmic solutions (in terms of efficiency), and yet are apparently straightforward for human chemists: detecting the overall *shape* and *class* of a presented molecule, in molecules that are large and composed of relatively homogeneous parts interconnected in cycles (such as the class of fullerene molecules [23]). As discussed in [18], determination of overall shape and chemical class for these classes of molecules is particularly challenging since the dense interconnection of the atoms in multiple fused cycles and the homogeneity of the atom environments. We hypothesise that a contributing factor in this performance discrepancy is that the use of a visual language in chemistry enables humans to directly harness the ‘gestalt’ or shape-detecting features of their visual perceptual machinery, *seeing the whole molecule at once* through the diagrammatic depiction, and therefore not needing to do the same sorts of computations that our algorithms need to do. If this line of thinking is correct, we should observe that the ability of humans to perform these tasks is affected by perturbances in the diagrammatic depiction more than in the size of the molecule. We conducted a study in which we time experi-

---

<sup>2</sup>As discussed in [18], we ignore statistical ‘black box’ approaches since they do not allow for explanations of their deductions and are not provably correct.

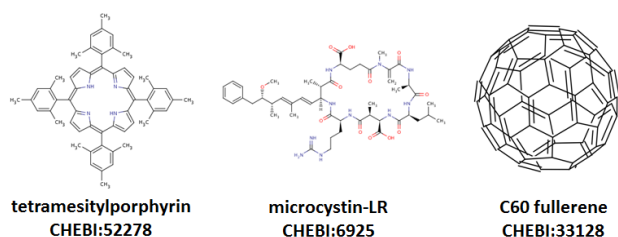
enced chemists performing a classification task on molecular diagrams with varied (a) diagrammatic faithfulness, and (b) size of the chemicals. We then evaluated the speed and accuracy of the chemists’ performance given these variances.

The remainder of this document is organised as follows. In the next section, we give our experimental design in the context of some background information about chemical diagrams and the class of chemical problems that we will use as a case study. Thereafter, we present our results and discussion. With only three participants and only 30 diagrams included in our experiment, our results can be considered a pilot study rather than a conclusive investigation. However, we consider these preliminary findings suggestive of future research directions, and we go on to further speculate about the implications for the use of artificial intelligence in chemistry applications.

## 1. Methods

### 1.1. The diagrammatic language of chemistry

Molecular entities are commonly represented visually as connected graphs, in which the vertices represent atoms (or groups) and the edges chemical bonds [39]. Cheminformatics software use the underlying graph as chemical data structure that serves as input to algorithmic calculations of features of the chemicals. Logic-based approaches also use a graph-based underlying representation as input to automated reasoning processes [26]. For human consumption, however, the underlying graph is projected onto a two-dimensional plane for visual interpretation. This diagrammatic depiction is a core offering of almost all chemical databases, and professional chemists develop an aptitude at discerning molecular features via such representations. Some examples of chemical diagrams are illustrated in Figure 1.



**Figure 1.** Some examples of chemical diagrams.

Chemical diagrams serve an invaluable purpose for chemists: they enable rapid evaluation of the overall chemistry of a given molecule, detection of errors or problems in the chemical structure being represented (e.g. infeasibility or chemical instability), and assessment of the properties or classifications that are relevant for the given molecule.

Chemical diagrams, like maps, represent *spatial* information. We have earlier referred to such spatial representations such as street maps, chemical diagrams, and engineering design models as *structural diagrams* [12], and they were called *analogical representations* in [36]. Here, we will focus not on the structural associations that we

highlighted before, but on the features of the overall shape and layout of chemicals that are available in chemical diagrams.

Molecular flexibility is also very important for molecular shape [13], as one 2D depiction can, through flexibility, yield many different 3D conformers that have vastly different properties in vitro. Such flexibility is not explicitly represented in 2D illustrations of molecules, but can be inferred from such representations given appropriate chemical knowledge.

### 1.2. Chemical shape perception task

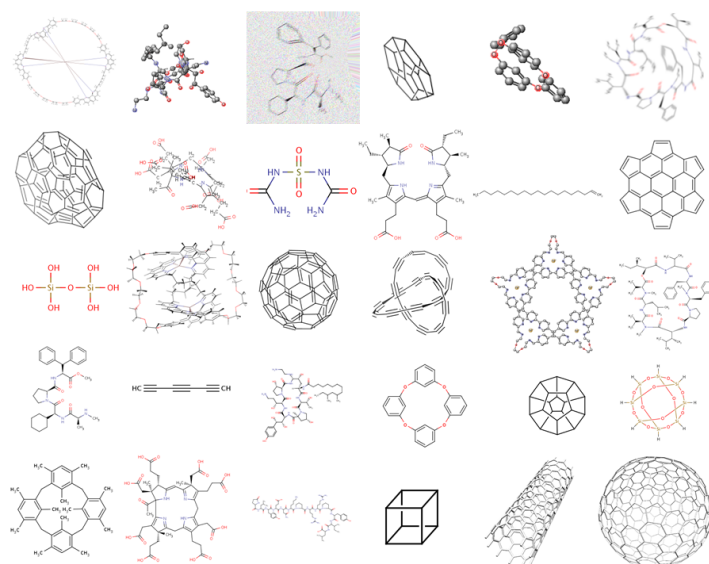
For our experiment, we have deliberately chosen classes of molecule that are known to be challenging to represent with logic-based automated reasoning approaches. Earlier, we have conducted an evaluation of the capabilities of algorithmic and logic-based approaches to reasoning tasks with molecular structures in [18]. The classes we selected for use in this task are:

1. Macrocyclic molecules, including calixarenes;
2. Polycyclic cages, including several differently sized fullerenes;
3. Shape-characterised molecules such as the catenanes and molecular knots;
4. Molecules that were not members of the above three classes as ‘controls’.

Macrocyclic molecules are molecules that form a large cyclic structure composed of linkages of smaller functional groups. Polycyclic cages are molecules that are composed entirely of cycles that are fused together in such a way as to form an overall cage-like structure, which is a feature that has interesting applications in medicinal chemistry and in materials science as the structure can serve to protect or capture a smaller molecule on the inside, or be engineered to lengthy tubes that are very strong. Examples are the fullerenes, cucurbiturils (named for their similarity to pumpkins), nanotubes, and small regular compounds such as cubane. Such nanomaterials have recently shown promise in the challenge of capturing highly volatile nerve agents and thereby preventing damage in vivo [22]. Molecules with specific shapes are of interest in the development of molecular machines, including the presence of stationary and movable parts, and the ability to respond with controlled movements to the external environment. Molecules that are mechanically interlocked—such as bistable rotaxanes and catenanes are some of the most intriguing systems in this area because of their capacity to respond to stimuli with controlled mechanical movements of one part of the molecule (e.g. one interlocked ring component) with respect to the other stationary part [10]. Similarly, molecules which display unusual energetic properties by virtue of their overall shape, such as molecular Möbius strips and trefoil knots, are an active research area for many novel applications, and in many cases mimic the extraordinary properties of biomolecular machinery such as active sites within protein complexes [34,42].

We selected five individual molecule types for the first two classes (macrocyclic molecules and cages). For the shape-characterised molecules, we were not able to find as many representatives in public chemical databases (our main source was the ChEBI database [15]), therefore we selected only four examples. Eight molecules that were not members of any of the three target classes but which were highly similar to one of the selected molecules (based on cheminformatics similarity scoring using Tanimoto over the molecular fingerprint, as implemented in OrChem [33]). Molecules were selected ranging from small to large, as measured in terms of counts of non-hydrogen atoms.

A randomly selected subset of eight of those 22 molecules was then subjected to diagrammatic distortion. Different distortion mechanisms were used. Firstly, the original molecule was computationally assigned a 3D conformation, which was then projected back onto a 2D diagram (a common outcome of computational processing of chemicals originally drawn by human chemists). Secondly, computational procedures for ‘clean’ 2D diagram generation were used. Finally, some of the diagrams were subjected to image processing to obscure the standard chemical representation either through blurring or shape-based transformation. The total number of diagrams was thus 30. The full set of molecules is shown in Figure 2.

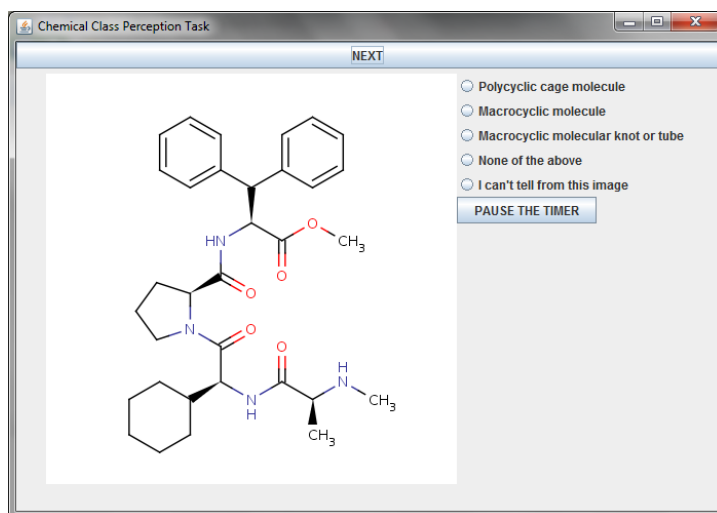


**Figure 2.** The molecule diagrams used in the experiment, including those showing distortions.

These diagrams were then displayed to the three experienced chemist participants in a random sequence<sup>3</sup>. For each diagram, the chemist was asked to determine the chemical class of the molecule, presented with the three classes, a fourth option ‘none of the above,’ and a final option ‘unable to tell from this diagram.’ Participants were timed as they completed the task, and their accuracy and agreement were calculated. Figure 3 shows a screenshot of the interface we developed in order to complete the perceptual task.

<sup>3</sup>There were three participants, each of whom had an academic background in chemistry and interacted with chemical data on a daily basis. The participants were explained the purpose of the experiment and each gave their informed consent. All data were stored anonymously and securely.





**Figure 3.** The chemical perception task interface, showing the chemical diagram and class selection options.

## 2. Results

### 2.1. The effect of image distortion on performance

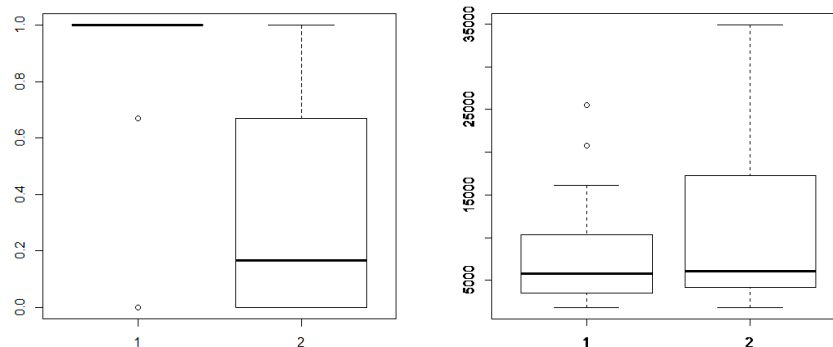
Image distortion had a significant effect on the accuracy of the chemical raters in choosing the correct classification for the classes. Figure 4 (a)<sup>4</sup> shows a boxplot of the classification task accuracy for the standard images as compared to the distorted images. The time taken (Figure 4 (b)) shows less of an effect than the accuracy, with the means not significantly different but the variance much larger in the case of the distorted images.

Ordinarily, chemists would look for additional information in case they encounter a partially obscured image and needed to determine the chemical class. Therefore, we do not restrict here our measure of accuracy to the percentage of correct classifications. Agreement between chemists in a classification task is an alternative measure of accuracy, which is especially useful in case the correct classification is not known in advance, but can supplement the known accuracy score used above with a clue as to the difficulty of the task. It might, for example, have been the case that the chemists had all agreed on incorrect classifications for the distorted images, leading to low accuracy but high agreement. However, agreement also differed strongly between the non-distorted and the distorted set of images, with the distorted images having a much lower agreement as measured by Cohen's Kappa statistic for multiple raters [6]. For the non-distorted pictures, the kappa was 0.88. For the distorted pictures, the kappa was 0.46%.

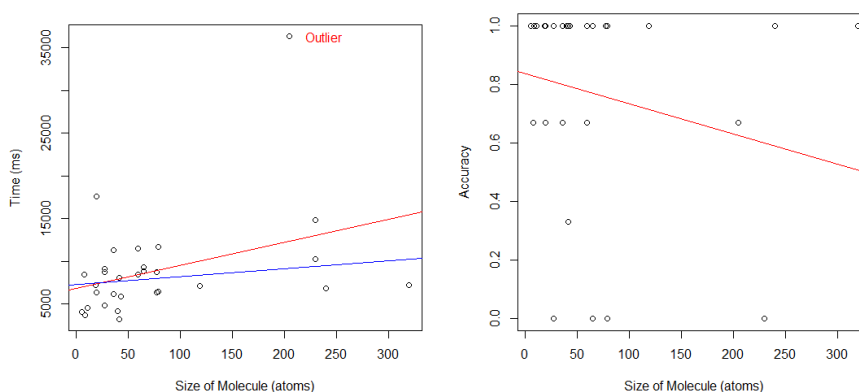
#### 2.1.1. The effect of size on performance

The scatter plot in Figure 5 (a) shows that size did not have a large impact on the time taken to perform the task. The red correlation line shows a weak positive correlation

<sup>4</sup>For space considerations, we do not present the full raw data result table here. However, this is available on request.



**Figure 4.** The display shows boxplots of (a) accuracy and (b) time taken (in ms), comparing the results for good vs. obscured visual layouts of the chemicals



**Figure 5.** Scatter plot of average (a) time to complete task (ms), and (b) accuracy, against size of the molecule.

between size and time taken to complete the task. However, this correlation is largely influenced by one data point, which itself depends on just one data point. The blue line shows the much weaker correlation that results from excluding the single outlier from the analysis.

Figure 5 (b) shows that accuracy was slightly anti-correlated with the size of the molecule, but this effect is not significant, with the p-value of the correlation only 0.24, and the 95% confidence interval for the correlation coefficient was from -0.54 to 0.15.

These results can be compared to algorithmic approaches and logic-based approaches for the relevant sort of feature detection that would be required to automatically compute the same task, i.e. automatic classification into the correct class based on chemical structures. Unfortunately, there is not yet any available generic system that is

able to perform the classifications tasks that were used in this experiment with which we may have compared the performance to our human experts. Indeed, as discussed in [18] our research has the long-term objective of enabling the development of just such a system, however, at this preliminary stage we do not have an available benchmark but must instead look to the performance profiles of algorithms that are known to be relevant.

## 2.2. Algorithmic cheminformatics approaches

The relevant algorithms that would be required to detect the classes specified include the detection of *subgraph isomorphism* and finding the *smallest set of rings* [18,43]. These algorithms are known to scale supralinearly in the number of atoms. For example, subgraph isomorphism in the general case is known to be NP-complete [7], although optimisations exist for various sub-classes of molecules, such as those that are planar [8].

For the particularly shape-defined classes, shape similarity algorithms on molecular structures exist that use ray-tracing of the projected surfaces of molecules to estimate the overall shape of the molecule and use that as a descriptor e.g. in virtual screening [2]. These methods depend on 3D conformer though, and for flexible molecules many conformers may result from the same 2D diagrammatic depiction, dramatically decreasing the performance of the algorithm.

Furthermore, a separate algorithm implementing a check on the rules of class membership would need to be hand-written for each of the three class types used in this task (macrocyclic, cage, shape-defined). This hampers the extensibility and flexibility of a system that needs to classify molecules in the general case [18,25]. On the other hand, logic-based systems address these objectives of being generic, extensible and flexible.

## 2.3. Logic-based approaches

The popular Web Ontology Language, OWL [11], is highly efficient in representing tree-like structures, but is unable to correctly represent cyclic structures [16]. A first-order logic programming based formalism has been proposed specifically for the case of representing chemical structures [26,25]. These description graph logic programs (DGLP) are able to represent objects whose parts are interconnected in arbitrary ways, including cyclic structures. The decidability of logic programs do not rely on the tree-model property that underlies the description logics behind OWL. However, representation of classes with more advanced overall topological features such as polycyclic cages is beyond the expressivity of DGLP as it requires quantification over all atoms in a molecule rather than specific atoms, parts or properties within the molecule.

Perhaps motivated by similar concerns on the limits of the logic-based approaches underlying languages such as OWL, Maojo *et al.* propose a ‘morphospacial’ approach to ontology with application in the nanomaterials domain [27]. Shape features are explicitly encoded in their ontology alongside other features such as composition. However, this approach merely pushes the problem onto those computational methods that are needed to derive the shape features automatically from some representation of the input chemical structure and thereby assign appropriate ontological categories to nanomolecular structures.

An approach for the representation of the overall structure or topology of highly symmetrical polycyclic molecules is described in [17,23]. There, the authors propose us-

ing a combination of monadic second-order logic and ordinary OWL, with a heterogeneous logical connection framework used to bridge between the two formalisms. This approach has not yet been implemented in practice, but shows promise for logical reasoning over features involving regularity in the overall structure of molecules. However, arbitrary entailment in monadic second-order logic is known to be computationally expensive<sup>5</sup>.

Spatial logics and spatial axiomatizations have been advanced in which it is possible to perform computational deductive reasoning [5,41]. However, it is not immediately straightforward to represent the problem of determining from an arbitrary chemical graph whether it is a member of the class of fullerenes (for example) as a spatial reasoning problem. We will develop this research question further in future work.

### 3. Discussion and Conclusions

While this study is small and exploratory in nature only, our results provide tentative support for a role for perception in human performance in the presented classification decision task, in that observed performance appeared to decrease with the quality of the diagrammatic representation rather than the size of the molecule. On the other hand, it is known that the best algorithmic and logical approaches to solving these particular tasks scale dramatically in the size of the molecule, rendering their habitual application to large numbers of molecules in a database problematic.

Larkin and Simon [24] attribute observed efficiencies of diagrammatic reasoning relative to non-diagrammatic reasoning to efficiencies in searching and inference in the representation space compared to that of a non-diagrammatic representation space, e.g. axioms. This may indeed be the root explanation, but it doesn't give guidance on how best to expose the representational efficiency that humans have (the ability to perceive the overall shape and connectivity in molecular diagrams) to computational processes. Traditional logical reasoning relies on linguistic or symbolic representation of the properties of objects together with the rules for deriving inferences on those properties. By contrast, diagrammatic representation can explicitly encode the relevant properties of objects and their background constraints such that the needed inferences can be directly drawn from the spatial constraints evident in the illustration [32], known as the "free ride" property.

Systems have been developed that enable the representation of logical axioms diagrammatically and the formalisation of accompanying reasoning systems to the extent that diagrammatic and traditional syllogistic reasoning can be combined in order to serve as an aid for human capability [29,28]. Such logical diagrammatic representations do not correspond directly to portions of reality, as the diagrammatic representations of chemicals correspond to classes of chemicals, but the correspondence is still analogical, i.e. by analogy. For example, Euler diagrams represent axioms such as *All A are B* as a smaller circle *A* entirely enclosed in a larger circle *B* [35]. This is analogous to *spatial inclusion*, as (for example) a smaller fullerene molecule can be fully enclosed in a larger fullerene molecule [23], and we could make corresponding statements such as *All atoms in molecule A are IN<sub>SPATIALLY</sub> molecule B*.

---

<sup>5</sup>Automated theorem provers such as LEO-II (<http://www.ags.uni-sb.de/~leo/>) are able to approximate some aspects of entailment checking.

Where perception is used as an aid to reasoning, care must be taken in the choice of the visual representation. For example, when visual diagrams are used as an aid to human logical reasoning, it has been found that Euler diagrams are more effective than Venn diagrams [28]. Irrelevant and distracting visual detail acts as a hindrance to reasoning rather than as an aid [21]. In the chemistry domain, for the class of classification problems we are interested in, this may be particularly important. Exposing the specifically visual information of a chemical diagram to computational processes would introduce additional constraints on the representation of the chemicals that currently only obtain in case the representation is intended for human consumption. Visual inference can sometimes be much more expensive than normal inference in the corresponding axiomatization, especially when the visual information is incomplete or, as we have tested, perturbed [1]. Adherence to standards for clear and unambiguous diagrammatic representation such as those put forward in [4] would go some way to address this concern in the chemistry domain.

In chemical similarity searching and bioactivity predictive modelling, quantitative shape-based 3D descriptors have met with mixed results stemming, on the one hand, from their greater computational cost than their 2D counterparts, and on the other hand, from the additional ‘noise’ that they can introduce in flexible molecules due to the variety of conformations [40]. One direction for our future research will be to evaluate the performance of these shape-based descriptors in assigning shape classes to molecules, such as ‘spherical’ and ‘cubic’. We are not aware of any existing work that applies this type of descriptor to the problem of structure-based chemical classification.

Our result emphasises the need for hybrid reasoning systems in chemistry that are able to combine features derived diagrammatically from visual representations of the molecule with the now-standard logic-based and algorithmic reasoning over the graph-based structure. Such hybrid systems have been advanced in other domains. For example, the Vivid system offers some diagrammatic reasoning capability alongside logical reasoning capability [1]. However, this system depends on algorithmic processes that “observe” pre-defined features in the diagrams included in the system capability. In the case of the chemical diagrams that form our case study here, some features are features of the *whole* diagram for which computational “observation” algorithms do not (to the best of our knowledge) yet exist. Research in machine vision may yield some methods that can be harnessed in pursuit of this objective [38].

Sloman [37] speculates that the ability to integratively process different types of representation with correspondingly different reasoning tasks might be a distinctive feature of intelligence in general; it is certainly a feature of human intelligence.

#### *Acknowledgements*

JH thanks the Swiss Center for Affective Sciences, and the European Commission via EU-OPENSREEN, for funding.

#### **References**

- [1] Arkoudas, K., Bringsjord, S.: Vivid: A framework for heterogeneous problem solving. Artificial Intelligence (Jun 2009), <http://dx.doi.org/10.1016/j.artint.2009.06.002>

- [2] Ballester, P.J., Richards, W.G.: Ultrafast shape recognition for similarity search in molecular databases. *Proc. R. Soc. A* 463, 1307–1321 (2007)
- [3] Berger, F., Flamm, C., Gleiss, P.M., Leydold, J., Stadler, P.F.: Counterexamples in chemical ring perception. *J Chem Inf Comput Sci* 44, 323–331 (2004)
- [4] Brecher, J., Degtyarenko, K.N., Gottlieb, H., Hartshorn, R.M., Hellwich, K.H., Kahovec, J., Moss, G.P., McNaught, A., Nyitrai, J., Powell, W., Smith, A., Taylor, K., Williams, A., Yerin, A., Town, W.: Graphical representation standards for chemical structure diagrams (IUPAC recommendations 2008). *Pure and Applied Chemistry* 80, 277–410 (2008)
- [5] Cabedo, L.M., Escrig, M.T.: A qualitative theory for shape representation and matching for design. In: R. Lopez de Mantaras and L. Saitta, editors, *Proceedings of the 16th European Conference on Artificial Intelligence* (2004)
- [6] Conger, A.: Integration and generalisation of kappas for multiple raters. *Psychological Bulletin* 88, 322–328 (1980)
- [7] Cook, S.A.: The complexity of theorem-proving procedures. In: *Proc. 3rd ACM Symposium on Theory of Computing*. p. 151158 (1971)
- [8] Eppstein, D.: Subgraph isomorphism in planar graphs and related problems. *Journal of Graph Algorithms and Applications* 3, 1–27 (1999)
- [9] Ertl, P., Rohde, B.: The molecule cloud - compact visualization of large collections of molecules. *Journal of Cheminformatics* 4(1), 12 (2012), <http://www.jcheminf.com/content/4/1/12>
- [10] Forgan, R.S., Sauvage, J., Stoddart, J.F.: Chemical topology: Complex molecular knots, links, and entanglements. *Chemical Reviews* 111, 5434–5464 (2011)
- [11] Grau, B.C., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P., Sattler, U.: OWL 2: The next step for OWL. *Web Semantics* 6, 309–322 (November 2008), <http://portal.acm.org/citation.cfm?id=1464505.1464604>
- [12] Hastings, J., Batchelor, C., Neuhaus, F., Steinbeck, C.: What's in an 'is about' link? Chemical diagrams and the IAO. In: *Proceedings of the International Conference on Biomedical Ontology (ICBO2011)*, Buffalo, USA (2011)
- [13] Hastings, J., Batchelor, C., Schulz, S.: Parts and wholes, shapes and holes in living beings. In: *Proceedings of the SHAPES 1.0 workshop*, Karlsruhe, Germany. CEUR-WS volume 812. (2011)
- [14] Hastings, J., Batchelor, C., Schulz, S., Jansen, L.: Collective bio-molecular processes: The hidden ontology of systems biology. *Proceedings of the UMoCoP workshop*, July 2012, Birmingham UK (2012)
- [15] Hastings, J., de Matos, P., Dekker, A., Ennis, M., Harsha, B., Kale, N., Muthukrishnan, V., Owen, G., Turner, S., Williams, M., Steinbeck, C.: The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic Acids Research Database issue*, gks1146 (2013)
- [16] Hastings, J., Dumontier, M., Hull, D., Horridge, M., Steinbeck, C., Sattler, U., Stevens, R., Hörne, T., Britz, K.: Representing chemicals using OWL, description graphs and rules. In: *Proc. of OWL: Experiences and Directions (OWLED 2010)* (2010)
- [17] Hastings, J., Kutz, O., Mossakowski, T.: How to model the shapes of molecules? combining topology and ontology using heterogeneous specifications. In: *Proceedings of the DKR Challenge Workshop*, Banff, Alberta, Canada, June 2011. (2011)
- [18] Hastings, J., Magka, D., Batchelor, C., Duan, L., Stevens, R., Ennis, M., Steinbeck, C.: Structure-based classification and ontology in chemistry. *Journal of Cheminformatics* 4(1), 8 (2012), <http://www.jcheminf.com/content/4/1/8>
- [19] Hey, A.J.G., Tansley, D.S.W., Tolle, K.M.: *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, Redmond, WA, USA (2009)
- [20] Kim, S., Bolton, E., Bryant, S.: PubChem3D: Shape compatibility filtering using molecular shape quadrupoles. *Journal of Cheminformatics* 3(1), 25 (2011), <http://www.jcheminf.com/content/3/1/25>
- [21] Knauff, M., Johnson-Laird, P.N.: Visual imagery can impede reasoning. *Memory and Cognition* 30, 363–371 (2002)
- [22] Kowalczyk, P., Gauden, P.A., Terzyk, A.P., Neimark, A.V.: Screening of carbonaceous nanoporous materials for capture of nerve agents. *Phys. Chem. Chem. Phys.* pp. – (2013), <http://dx.doi.org/10.1039/C2CP43366D>
- [23] Kutz, O., Hastings, J., Mossakowski, T.: Modelling Highly Symmetrical Molecules: Linking Ontologies and Graphs Artificial Intelligence: Methodology, Systems, and Applications. *Lecture Notes in Computer Science*, vol. 7557, chap. 11, pp. 103–111. Springer Berlin / Heidelberg, Berlin, Heidelberg (2012),

[http://dx.doi.org/10.1007/978-3-642-33185-5\\_11](http://dx.doi.org/10.1007/978-3-642-33185-5_11)

- [24] Larkin, J., Simon, H.: Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science* 11, 65–99 (1987)
- [25] Magka, D.: Ontology-based classification of molecules: A logic programming approach. In: *Proceedings of the SWAT4LS conference*, 30 November 2012, Paris, France. p. . (2012)
- [26] Magka, D., Motik, B., Horrocks, I.: Modelling structured domains using description graphs and logic programming. Tech. rep., Department of Computer Science, University of Oxford (2011)
- [27] Maojo, V., Fritts, M., Martín-Sánchez, F., de la Iglesia, D., Cachau, R.E., García-Remesal, M., Crespo, J., Mitchell, J.A., Anguita, A., Baker, N., Barreiro, J.M., Benitez, S.E., de la Calle, G., Facelli, J.C., Ghazal, P., Geissbühler, A., Gonzalez-Nilo, F., Graf, N.M., Grangeat, P., Hermosilla, I., Hussein, R., Kern, J., Koch, S., Legré, Y., López-Alonso, V., López-Campos, G., Milanesi, L., Moustakis, V., Munteanu, C.R., Otero, P., Pazos, A., Pérez-Rey, D., Potamias, G., Sanz, F., Kulikowski, C.A.: Nanoinformatics: developing new computing applications for nanomedicine. *Computing* 94(6), 521–539 (2012)
- [28] Mineshima, K., Okada, M., Takemura, R.: Two types of diagrammatic inference systems: Natural deduction style and resolution style. In: *Diagrams 2010, Lecture Notes In Artificial Intelligence* 6170. Springer (2010)
- [29] Mineshima, K., Okada, M., Takemura, R.: A generalized syllogistic inference system based on inclusion and exclusion relations. *Studia Logica* (2012)
- [30] P. V., J. F., N. R.: Understanding visualization: A formal approach using category theory and semiotics. *IEEE Trans Vis Comput Graph.* Sep 21. (2012)
- [31] Raymond, J.W., Willett, P.: Maximum common subgraph isomorphism algorithms for the matching of chemical structures. *Journal of Computer-Aided Molecular Design* 16, 521–533 (2002)
- [32] Recanatì, C.: Hybrid reasoning and the future of iconic representations. In: *Proceedings of the 2008 conference on Artificial General Intelligence 2008: Proceedings of the First AGI Conference*. pp. 299–310. IOS Press, Amsterdam, The Netherlands, The Netherlands (2008), <http://dl.acm.org/citation.cfm?id=1566174.1566202>
- [33] Rijnbeek, M., Steinbeck, C.: OrChem - An open source chemistry search engine for Oracle(R). *Journal of Cheminformatics* 1(1), 17 (2009), <http://dx.doi.org/10.1186/1758-2946-1-17>
- [34] Rzepa, H.S.: Wormholes in chemical space connecting torus knot and torus link  $\pi$ -electron density topologies. *Phys. Chem. Chem. Phys.* pp. 1340–1345 (2009)
- [35] Sato, Y., Mineshima, K., Takemura, R., M.Okada: On the cognitive efficacy of euler diagrams in syllogistic reasoning: A relational perspective. In: *Euler Diagram Workshop 2012, CEUR* 854 (2012)
- [36] Sloman, A.: Interactions between philosophy and ai: The role of intuition and non-logical reasoning in intelligence. In: *Proceedings of the Second International Joint Conference on Artificial Intelligence*. p. . (1971)
- [37] Sloman, A.: Musings on the roles of logical and non-logical representations in intelligence. In: *Glasgow, J., Narayanan, H., Chandrasekaran (eds.) Diagrammatic Reasoning: Computational and Cognitive Perspectives*, p. . AAAI Press (1995)
- [38] Sun, M., Bradski, G., Xu, B.X., Savarese, S.: Depth-encoded hough voting for coherent object detection, pose estimation, and shape recovery. In: *ECCV* (2010)
- [39] Trinajstić, N.: *Chemical graph theory*. CRC Press, Florida, USA (1992)
- [40] Venkatraman, V., Chakravarthy, P., Kihara, D.: Application of 3d zernike descriptors to shape-based ligand similarity searching. *Journal of Cheminformatics* 1(1), 19 (2009), <http://www.jcheminf.com/content/1/1/19>
- [41] Vincent Dugat, P.G., Larvor, Y.: Qualitative geometry for shape recognition. *Applied Intelligence* 17, 253–263 (2002)
- [42] Wannere, C.S., Rzepa, H.S., Rinderspacher, B.C., Paul, A., III, H.F.S., v. R. Schleyer, P., Allan, C.S.M.: The geometry and electronic topology of higher-order Möbius charged annulenes. *J. Phys. Chem.* 113, 11619–11629 (2009)
- [43] Wegner, J.K., Sterling, A., Guha, R., Bender, A., Faulon, J.L., Hastings, J., O’Boyle, N., Overington, J., Van Vlijmen, H., Willighagen, E.: *Cheminformatics. Communications of the ACM* 55(11), 65–75 (2012)

# Shapes as property restrictions and property-based similarity

Silvia LIKAVEC <sup>1</sup>

*Dipartimento di Informatica, Università di Torino, Italy*

**Abstract.** In this work we look into the details of modeling shapes in an ontology as property restrictions on classes. In this way shapes do not have to be categorized in an exhaustive hierarchy and there is no need to take immediate decisions on how to group the objects, it is rather possible to individualize some important characteristics of shapes and use them as the basis for their categorization and comparison. This approach also makes it possible to use the adequate similarity measure based on properties which helps find similar shapes in different contexts, depending on the relevance of the properties in the particular situation.

**Keywords.** ontology, OWL, shapes, properties, restrictions, similarity

## Introduction

In many different areas, from virtual reality to architectural design, from biology to medicine, from mathematics to computer science, the notion of shape plays a crucial role. Finding patterns and forms in objects surrounding us, describing them and understanding their interaction is essential to human nature. Hence, varied approaches to the categorization of shapes and forms, as well as their mutual similarity and connectedness, are of great importance for the development of many scientific fields.

In various domains, there is a raising tendency to use ontologies as powerful formalisms for knowledge representation with associated reasoning mechanisms (inheritance, subsumption, classification etc.). Ontologies provide explicit specifications of domain concepts and relationships that exist between them [1]. They guarantee exact semantics for each statement and avoid semantic ambiguities. Their usage allows for extensibility and re-usability, since they are expressed with standard formats and technologies.

In the domain of shape, form and structure representation, there were some attempts at modeling shapes ontologically, as an exhaustive hierarchy. In [9] the authors develop a limited graphics ontology for natural language interfaces, covering the concepts like “Shape”, “Action” and various features which describe shapes (size, color, position etc.). In [5] two first-order ontologies for representing 2D and 3D shapes like surfaces and boxes are introduced. They use only the notions such as part-of and connectedness, rather than Euclidean geometric relations, such as alignment and length of segments, or the

---

<sup>1</sup>Corresponding Author: Silvia Likavec, Dipartimento di Informatica, Corso Svizzera 185, 10149 Torino, Italy; E-mail: likavec@di.unito.it.



notions of curvature or surface area. In the domain of architectural design, [8] presents a conceptual “building shape ontology” which sorts building shapes and captures their meaning and semantics.

In the realm of spatial design and reasoning, authors in [2] explore the role ontological formalization plays in modeling of high-level conceptual requirement constraints. They concentrate on ontological modeling of structural forms from different perspectives. As for architectural design, information that is being used often originates from various sources. In [6], the authors take a step towards integration of various aspects of architectural domain (spatial constraints, relations among objects, abstract conceptualizations) designing modular ontologies based on the theory of  $\varepsilon$ -connections. [13] describes a method for the retrieval of 3-dimensional shapes (in this case furniture models) based on a mapping between low level features described by the shape descriptor and ontology concepts. This furniture ontology is used in annotation and key word based retrieval of furniture models.

As far as similarity among ontological concepts is concerned, three main approaches can be distinguished. The first one [12] is based on information content of a class in an IS-A taxonomy, given by the negative logarithm of the probability of occurrence of the class in a text corpus. The second approach [11] uses the ontology graph structure, by measuring directly the distance between nodes (usually, the number of edges or the number of nodes that need to be traversed in order to reach one node from the other). Finally, the third approach combines the information content approach with edge counting based approach (see for example [7]). Different notions of similarity and the relationships among them are tackled in [3]. Starting from Leibnizian relative identity as the only local form of similarity, they show that more sophisticated notions can be obtained by applying transformations across heterogeneous logics. They also distinguish between ontological and epistemic similarities. While ontological similarities stem from the structure of the world itself, epistemic similarities are used to connect entities in different worlds.

The rest of the paper is organized as follows. Section 1 provides a summary of the treatment of properties in OWL and the definition of property restrictions on classes. In Section 2 we discuss some issues concerning modeling of shapes as property restrictions on classes, followed by some examples of shape definitions. Details of the approach to calculating similarity of shapes based on properties can be found in Section 3. Section 4 concludes.

## 1. Properties and property restrictions in OWL

### 1.1. Properties in OWL

In different contexts, domain knowledge can be represented semantically using ontologies expressed in OWL<sup>2</sup>. In an ontology, domain concepts are organized hierarchically and have their features defined as properties. Two kinds of properties can be distinguished in OWL: (i) *object properties* relating individuals among themselves and (ii) *data type properties* relating individuals to data type values.

Characteristics of a property are defined with a property axiom, most commonly defining its domain and range. `rdfs:domain` links a property to a class description,

---

<sup>2</sup><http://www.w3.org/TR/owl-ref>

whereas `rdfs:range` links a property to either a class description or a data range. For example:

```
<owl:ObjectProperty rdf:ID="has_curvature">
  <rdfs:domain rdf:resource="#Shape"/>
  <rdfs:range rdf:resource="#Curvature" />
</owl:ObjectProperty>
```

defines the property *has\_curvature* which ties the elements of *Shape* class to the elements of *Curvature* class.

Equivalent properties are defined with `owl:equivalentProperty`.

In the following section we will see how properties are used in OWL to define classes with property restrictions.

### 1.2. Defining classes with property restrictions

Properties are used in OWL for defining classes with property restrictions, by means of *local anonymous classes*, which are collections of individuals all satisfying certain restrictions on certain properties. Two kinds of property restrictions exist: *value constraints* and *cardinality constraints*. A value constraint concerns constraints on the range of the property when applied to a particular class description. A cardinality constraint imposes constraints on the number of values a property can take, in the context of a particular class description.

We start with the brief description of value constraints. There are three ways of defining value constraints:

- `owl:allValuesFrom` defines a class for which all the values of the given property are either members of the specified class or data values within the specified data range. It is possible not to have any values for the given property. For example:

```
<owl:Restriction>
  <owl:onProperty rdf:resource="#has_angle" />
  <owl:allValuesFrom rdf:resource="#RightAngle" />
</owl:Restriction>
```

describes an anonymous OWL class of all individuals for which the *has\_angle* property only has values of the class *RightAngle* (for example square or rectangle). In predicate logic, the counterpart of `owl:allValuesFrom` constraint is the universal quantifier, i.e. for each instance of the class defined with the restriction, every value for the property must fulfill the constraint and the constraint is trivially satisfied for an instance that has no value for the specified property.

- `owl:someValuesFrom` specifies a class for which at least one of the values of the given property is either a member of the specified class or a data value within the specified data range (at least one must exist). There might be other values for the given property. For example:

```
<owl:Restriction>
  <owl:onProperty rdf:resource="#has_angle" />
  <owl:someValuesFrom rdf:resource="#RightAngle" />
</owl:Restriction>
```

describes an anonymous OWL class of all individuals which have at least one right angle (for example right-angled triangle). In predicate logic, the counterpart of `owl:someValuesFrom` constraint is the existential quantifier, i.e. for each instance of the class defined with the restriction, there exists at least one value for the property that fulfills the constraint.

- `owl:hasValue` defines a class for which the specified property has at least one value semantically equal to the specified value, which can be either an individual or a data value<sup>3</sup>. For example:

```
<owl:Restriction>
  <owl:onProperty rdf:resource="#contains" />
  <owl:hasValue rdf:resource="#Circle100" />
</owl:Restriction>
```

describes an anonymous OWL class which contains a specific circle.

On the other hand, cardinality constraints can be expressed by using one of the following three constraints:

- `owl:maxCardinality` describes a class with at most *max* semantically distinct values for the specified property (*max* being the value of the cardinality constraint). For example

```
<owl:Restriction>
  <owl:onProperty rdf:resource="#has_number_of_edges" />
  <owl:maxCardinality rdf:datatype="&xsd;nonNegativeInteger">5
</owl:maxCardinality>
</owl:Restriction>
```

describes an anonymous OWL class of individuals that have at most five edges (for example polygons with 3 or 4 or 5 edges),.

- `owl:minCardinality` is defined analogously to `maxCardinality`, where the class has at least *min* semantically distinct values.
- `owl:cardinality` is defined analogously to `maxCardinality`, where the class has exactly *m* semantically distinct values. It is actually a redundant concept, since it can be defined with the combination of `maxCardinality` and `minCardinality`.

From the above we can see that we can consider each of the concepts in our ontology, to have certain properties defined for it. These properties further describe the concepts in the ontology and can be used to categorize them and to calculate their mutual similarity.

### 1.3. Instances and their properties

An instance in the ontology is defined with individual axioms called “facts” which describe its class membership, property values and individual identity. An instance is related to the class it belongs to directly with the `rdfs:type` relation and basically inherits

---

<sup>3</sup>For datatypes “semantically equal” means that the lexical representation of the literals maps to the same value. For individuals it means that they either have the same URI reference or are defined as being the same individual with `owl:sameAs`.

the properties of the class it belongs to. Hence, in OWL the properties of the instances are defined by associating to each property its specific value. For example, the following describes a red circle with the radius equal to 3 and a dotted outline.

```
<Circle rdf:ID="Circle100">
  <has_radius rdf:datatype="xsd:positiveInteger">3</has_radius>
  <has_outline rdf:resource="#Dotted"/>
  <has_color rdf:resource="#Red"/>
</Circle>
```

## 2. Shapes defined as class restrictions

Categorization of the world around us is inherent to human perception and reasoning. Many objects are internalized easier if they are reduced to simpler forms and shapes that we are familiar with and that we can easily group with other similar objects.

We give some directions on how to model two-dimensional shapes, since they are the easiest to comprehend and understand.<sup>4</sup> The most common categorization of two-dimensional shapes is according to the kind of edges the shapes contain to curved shapes and straight line composed shapes (polygons). But another categorization could start from convex and concave shapes. Or we might want to categorize the shapes based on the number of edges they have. The possibilities are many.

So instead of forcing this somehow artificial categorization upon the shape world, we would do the shapes more justice by defining them as property restrictions on classes. We can start by defining many different properties which would help us precisely describe the shapes we need. For example, the property *has\_edge\_kind* could be used to define as property restrictions the classes *CurvedShape* and *Polygon* (shapes composed from straight lines), whereas the property *has\_curvature* would be used to distinguish *ConvexShape* class from *ConcaveShape* class (again defining them as restrictions). In this way, there is no need to a-priori decide which categorization should happen higher up in the hierarchy, they can peacefully co-exist together (and not be the only ones). Once the first level is modeled, we can proceed to model their subclasses. At this point we can have direct subclasses with additional properties or additional restrictions. So we can include properties like *number\_of\_edges*, *number\_of\_equal\_edges*, *number\_of\_parallel\_edges* etc. This would also help us compare the shapes having all these properties defined explicitly for them.

In this light, let us have a look at two shape definitions, namely rhombus and rectangle. A rhombus can be defined as a simple (non-self-intersecting) quadrilateral with four equal edges, whereas a rectangle can be defined as quadrilateral with four right angles (and they are both convex). So if we define the *Quadrilateral* class as a subclass of *Polygon* class which has the property *has\_number\_of\_sides* restricted to 4, we can define rhombus and rectangle as follows:

```
<owl:Class rdf:ID="Rectangle">
  <rdfs:subClassOf rdf:resource="#Quadrilateral"/>
  <rdfs:subClassOf rdf:resource="#ConvexShape"/>
```

---

<sup>4</sup>Three-dimensional and n-dimensional objects are treated similarly.

```

<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty rdf:resource="#has_angle" />
    <owl:allValuesFrom rdf:resource="#RightAngle" />
  </owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty rdf:resource="#has_number_of_angles"/>
    <owl:cardinality rdf:datatype="&xsd;nonNegativeInteger">4
  </owl:cardinality>
  </owl:Restriction>
</rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:ID="Rhombus">
  <rdfs:subClassOf rdf:resource="#Quadrilateral"/>
  <rdfs:subClassOf rdf:resource="#Convex"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#has_number_of_equal_edges"/>
      <owl:cardinality rdf:datatype="&xsd;nonNegativeInteger">4
    </owl:cardinality>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

```

Obviously, these are not the only ways to define these two, or any other shape. The process is versatile and applicable in many different contexts. Above all, it enables very natural comparison of shapes and establishes their similarity based on properties, as we will see in the following section.

### 3. Property-based similarity of shapes

If we define shapes as property restrictions (on values and cardinality), we can find similar shapes by comparing their properties. The property-based similarity of two shapes  $S_1$  and  $S_2$ , can be calculated by starting from Tversky's feature-based model of similarity [14], where similarity between objects is a function of both their common and distinctive features:

$$\text{sim}_T(S_1, S_2) = \frac{\alpha(\psi(S_1) \cap \psi(S_2))}{\beta(\psi(S_1) \setminus \psi(S_2)) + \gamma(\psi(S_2) \setminus \psi(S_1)) + \alpha(\psi(S_1) \cap \psi(S_2))}. \quad (1)$$

Here  $\psi(S)$  is the function which describes all the relevant features of  $S$ , and  $\alpha, \beta, \gamma \in \mathbb{R}$  are parameters which permit us to treat differently the various components. For  $\alpha = 1$  maximal importance is assigned to the common features of the two shapes and for  $\beta = \gamma$  non-directional similarity measure is achieved. We will use  $\alpha = \beta = \gamma = 1$ .

Hence, to be able to use Tversky's model we need to calculate the following:

- *common features of  $S_1$  and  $S_2$* :  $CF(S_1, S_2) = \psi(S_1) \cap \psi(S_2)$ ,
- *distinctive features of  $S_1$* :  $DF(S_1) = \psi(S_1) \setminus \psi(S_2)$  and
- *distinctive features of  $S_2$* :  $DF(S_2) = \psi(S_2) \setminus \psi(S_1)$ .

Putting these values into the formula (1) and taking  $\alpha = \beta = \gamma = 1$  we obtain:

$$SIM_T(S_1, S_2) = \frac{CF(S_1, S_2)}{DF(S_1) + DF(S_2) + CF(S_1, S_2)}. \quad (2)$$

In order to calculate common and distinctive features for  $S_1$  and  $S_2$ , for each property  $p$ , we calculate  $CF_p$ ,  $DF_p^1$  and  $DF_p^2$ , which denote how much the property  $p$  contributes to common features of  $S_1$  and  $S_2$ , distinctive features of  $S_1$  and distinctive features of  $S_2$ , respectively. In what follows we would see how different ways of defining properties in OWL influence the calculation of these values. We consider equal the properties defined with `owl:EquivalentProperty`.

We start from three kinds of value restriction declarations: (i) `owl:allValuesFrom`; (ii) `owl:someValuesFrom`; (iii) `owl:hasValue`. Based on how the restrictions on properties are defined for  $S_1$  and  $S_2$ , we can distinguish the following six cases:

1. The property  $p$  is defined with `owl:allValuesFrom` for both  $S_1$  and  $S_2$ . Let the property  $p$  be defined for  $S_1$  with  
 $\langle \text{owl:allValuesFrom rdf:resource} = \text{"\#A_1"} \rangle$   
and for  $S_2$  with  
 $\langle \text{owl:allValuesFrom rdf:resource} = \text{"\#A_2"} \rangle$ .  
Let  $a_1$  (resp.  $a_2$ ) be the number of sub-classes of  $A_1$  (resp.  $A_2$ ). If  $A_1$  and  $A_2$  are equal classes or declared equivalent with `owl:equivalentClass` or  $A_1$  is a subclass of  $A_2$ , then  $CF_p = \frac{1}{(a_1 + 1)}$ . Otherwise  $DF_p^1 = \frac{1}{a_1 + 1}$  and  $DF_p^2 = \frac{1}{a_2 + 1}$ .
2. The property  $q$  is defined with `owl:someValuesFrom` both for  $S_1$  and  $S_2$ . Let the property  $q$  be defined for  $S_1$  with  
 $\langle \text{owl:someValuesFrom rdf:resource} = \text{"\#B_1"} \rangle$   
and for  $S_2$  with  
 $\langle \text{owl:someValuesFrom rdf:resource} = \text{"\#B_2"} \rangle$ .  
Let  $b_1$  (resp.  $b_2$ ) be the number of sub-classes of  $B_1$  (resp.  $B_2$ ) and  $w$  be the number of classes in the whole domain. If  $B_1$  and  $B_2$  are equal classes or declared equivalent with `owl:equivalentClass` or  $B_1$  is a subclass of  $B_2$ , then  $CF_q = \frac{1}{(b_1 + 1)w}$ . Otherwise  $DF_q^1 = \frac{1}{(b_1 + 1)w}$  and  $DF_q^2 = \frac{1}{(b_2 + 1)w}$ .
3. Let property  $r$  be defined for  $S_1$  with  
 $\langle \text{owl:hasValue rdf:resource} = \text{"\#V_1"} \rangle$   
and for  $S_2$  with  
 $\langle \text{owl:hasValue rdf:resource} = \text{"\#V_2"} \rangle$   
If  $V_1$  and  $V_2$  are the same values or declared same with `owl:sameAs`, then  $CF_q = 1$ . Otherwise  $DF_q^1 = 1$  and  $DF_q^2 = 1$ .
4. If the property  $t$  is defined for  $S_1$  with  
 $\langle \text{owl:hasValue rdf:resource} = \text{"\#V_3"} \rangle$   
and for  $S_2$  with

$\langle \text{owl:allValuesFrom rdf:resource}=\text{"\#A}_3\text{"} \rangle$ ,

and if  $V_3$  is an instance of  $A_3$  or one of its subclasses, then  $cf_t = 1$ . Otherwise, if  $a_3$  is the number of sub-classes of  $A_3$ , then  $df_t^1 = 1$  and  $df_t^2 = \frac{1}{a_3 + 1}$ .

5. If the property  $x$  is defined for  $S_1$  with

$\langle \text{owl:hasValue rdf:resource}=\text{"\#V}_4\text{"} \rangle$

and for  $S_2$  with

$\langle \text{owl:someValuesFrom rdf:resource}=\text{"\#B}_3\text{"} \rangle$

and if  $V_4$  is an instance of  $B_3$  or one of its subclasses, then  $cf_x = 1$ . Otherwise, if  $b_3$  is the number of sub-classes of  $B_3$  and if  $w$  is the number of classes in the whole domain, then  $df_x^1 = 1$  and  $df_x^2 = \frac{1}{(b_3 + 1)w}$ .

6. If the property  $y$  is defined for  $S_1$  with

$\langle \text{owl:allValuesFrom rdf:resource}=\text{"\#A}_4\text{"} \rangle$

and for  $S_2$  with

$\langle \text{owl:someValuesFrom rdf:resource}=\text{"\#B}_4\text{"} \rangle$

and if  $a_4$  (resp.  $b_4$ ) is the number of sub-classes of  $A_4$  (resp.  $B_4$ ) and  $w$  is the number of classes in the whole domain, then  $cf_y = \frac{1}{(a_4 + 1)(b_4 + 1)w}$ ,  $df_y^1 = \frac{1}{a_4 + 1}$  and  $df_y^2 = \frac{1}{(b_4 + 1)w}$ .

Next we consider three kinds of cardinality restriction declarations: (i) `minCardinality`; (ii) `maxCardinality`; (iii) `cardinality`. We can distinguish the following cases:

1. If the property  $f$  is defined with `owl:maxCardinality` for both  $S_1$  and  $S_2$ , and it has value  $m$  in  $S_1$  and value  $n$  in  $S_2$ , where  $m \leq n$ , then  $cf_f = \frac{1}{n - 1}$ ,  $df_f^1 = 0$  and  $df_f^2 = n - m$ . The case when  $m \geq n$  is analogous.
2. If the property  $g$  is defined with `owl:minCardinality` for both  $S_1$  and  $S_2$ , the values for this restriction would not contribute to common and distinctive features, since each of these restrictions can have infinitely many values. It only contributes to similarity calculation if it is declared together with `owl:maxCardinality` restriction, which is then the following case.
3. If the property  $h$  is defined with `owl:cardinality` for both  $S_1$  and  $S_2$ , and it has value  $m$  in  $S_1$  and value  $n$  in  $S_2$ , then if  $m = n$   $cf_f = 1$ . Otherwise, if  $m < n$  then  $df_f^1 = 0$  and  $df_f^2 = n - m$ . The case when  $m > n$  is analogous.

Of course, the subclass relation should be taken into account, hence providing each class with the property definitions inherited from parent classes.

Finally, to calculate all common and distinctive features of  $S_1$  and  $S_2$  we repeat the above process for each property defined for  $S_1$  and  $S_2$ , obtaining:

$$\begin{aligned}
CF(S_1, S_2) &= \sum_{i_p=1}^{n_p} CF_{p_{i_p}} + \sum_{i_q=1}^{n_q} CF_{q_{i_q}} + \sum_{i_r=1}^{n_r} CF_{r_{i_r}} + \sum_{i_t=1}^{n_t} CF_{t_{i_t}} + \sum_{i_x=1}^{n_x} CF_{x_{i_x}} + \sum_{i_y=1}^{n_y} CF_{y_{i_y}} \\
&\quad + \sum_{i_f=1}^{n_f} CF_{f_{i_f}} + \sum_{i_g=1}^{n_g} CF_{g_{i_g}} + \sum_{i_h=1}^{n_h} CF_{h_{i_h}} \\
DF(S_1) &= \sum_{i_p=1}^{n_p} DF_{p_{i_p}}^1 + \sum_{i_q=1}^{n_q} DF_{q_{i_q}}^1 + \sum_{i_r=1}^{n_r} DF_{r_{i_r}}^1 + \sum_{i_t=1}^{n_t} DF_{t_{i_t}}^1 + \sum_{i_x=1}^{n_x} DF_{x_{i_x}}^1 \\
&\quad + \sum_{i_y=1}^{n_y} DF_{y_{i_y}}^1 + \sum_{i_f=1}^{n_f} DF_{f_{i_f}}^1 + \sum_{i_g=1}^{n_g} DF_{g_{i_g}}^1 + \sum_{i_h=1}^{n_h} DF_{h_{i_h}}^1 \\
DF(S_2) &= \sum_{i_p=1}^{n_p} DF_{p_{i_p}}^2 + \sum_{i_q=1}^{n_q} DF_{q_{i_q}}^2 + \sum_{i_r=1}^{n_r} DF_{r_{i_r}}^2 + \sum_{i_t=1}^{n_t} DF_{t_{i_t}}^2 + \sum_{i_x=1}^{n_x} DF_{x_{i_x}}^2 \\
&\quad + \sum_{i_y=1}^{n_y} DF_{y_{i_y}}^2 + \sum_{i_f=1}^{n_f} DF_{f_{i_f}}^2 + \sum_{i_g=1}^{n_g} DF_{g_{i_g}}^2 + \sum_{i_h=1}^{n_h} DF_{h_{i_h}}^2
\end{aligned}$$

where  $n_p$  (resp.  $n_q, n_r, n_t, n_x, n_y, n_f, n_g, n_h$ ) is the number of properties defined in each of the possible ways explained above. Finally, we calculate the similarity between two entities  $S_1$  and  $S_2$  defined with restrictions using the formula (2):

$$SIM(S_1, S_2) = \frac{CF(S_1, S_2)}{DF(S_1) + DF(S_2) + CF(S_1, S_2)}.$$

Another feature we want to take into account is the presence of equivalent classes, even though they are not defined as restrictions. We assume that two classes declared equivalent with `owl:equivalentClass` have similarity based on properties equal to 1.

As far as individuals are concerned (instances of the classes) we simply compare the property-value pairs for each instance. If the property  $p$  has  $h'$  different values in  $S_1$  and  $h''$  different values in  $S_2$ , and we denote by  $k$  the number of times  $S_1$  and  $S_2$  have the same value for  $p$ , then  $cf_p = \frac{k^2}{h'h''}$ ,  $df_p^1 = \frac{h' - k}{h'}$  and  $df_p^2 = \frac{h'' - k}{h''}$ . We repeat this for every property  $p_1, \dots, p_k$  used to describe the given instance. Finally, for instances of classes we obtain:

$$\begin{aligned}
CF(S_1, S_2) &= \sum_{i_{p_1}=1}^{n_{p_1}} CF_{p_{i_{p_1}}} + \dots + \sum_{i_{p_k}=1}^{n_{p_k}} CF_{p_{i_{p_k}}} \\
DF^1(S_1, S_2) &= \sum_{i_{p_1}=1}^{n_{p_1}} DF_{p_{i_{p_1}}}^1 + \dots + \sum_{i_{p_k}=1}^{n_{p_k}} DF_{p_{i_{p_k}}}^1 \\
DF^2(S_1, S_2) &= \sum_{i_{p_1}=1}^{n_{p_1}} DF_{p_{i_{p_1}}}^2 + \dots + \sum_{i_{p_k}=1}^{n_{p_k}} DF_{p_{i_{p_k}}}^2.
\end{aligned}$$

### 3.1. Relevance of properties

When defining a certain shape, not all the properties have the same importance in different contexts. For example, in one context two shapes would be regarded similar if they have similar number of angles and edges, in another one if they are of similar size or if they are both concave or convex. In the above presented approach, it is possible to account for relevance of properties by providing the *relevance factors*  $R_{i_p}^p$ ,  $i_p = 1, \dots, n_p$ , for each property  $p$ . Relevance factors can be either given as a-priori expert values or gathered as user preferences. In this way, some properties become more important than the others and the formula for calculating the mutual similarity between shapes  $S_1$  and  $S_2$  becomes:

$$SIM^r(S_1, S_2) = \frac{CF^r(S_1, S_2)}{DF^r(S_1) + DF^r(S_2) + CF^r(S_1, S_2)}$$



where  $\text{cf}^r(S_1, S_2) = \sum_{i_p=1}^{n_p} R_{i_p} \text{cf}_{p_{i_p}} + \dots + \sum_{i_h=1}^{n_h} R_{i_h} \text{cf}_{h_{i_h}}$  and similarly for  $\text{df}^r(S_1)$  and  $\text{df}^r(S_2)$ .

#### 4. Conclusions

When using ontologies to represent domain knowledge, not always it is convenient to represent shapes in an exhaustive hierarchy. It might be desirable to single out certain properties of shapes and then categorize them having these properties in mind. This is possible if shapes are defined as property restrictions on classes, both on value and cardinality. Representing shapes as property restrictions makes it possible to introduce a very natural similarity measure based on properties. This measure changes depending on the context in which it is being used, making it possible to give more relevance to certain properties in different situations. Apart from modeling shapes as property restrictions on classes, this approach would bring new insights into modeling forms and patterns as well, as it avoids strict categorizations, providing a flexible environment for expressing various features of complex forms.

The presented technique for calculating property-based similarity was first used in [4], for propagation of user interests in ontology based user model. It was evaluated in the context of PIEMONTE project [10]<sup>5</sup> which developed a framework based on intelligent objects composed from a real and a virtual part coexisting at the same time, in the context of gastronomy. Although this initial approach did not include cardinality restrictions it showed satisfying performance w.r.t. to actual reasoning and computation of similarity and helped improve the recommendation process. A future implementation of this method would include the cardinality restrictions and show how it handles them, providing feedback for any necessary adjustments.

#### References

- [1] G. Antoniou and F. van Harmelen, *A Semantic Web Primer*, second edition. The MIT Press, 2008.
- [2] M. Bhatt and J. Hois and O. Kutz and F. Dylla, *Modelling functional requirements in spatial design*. in *Proc. 29th international conference on Conceptual modeling, ER '10*, pp. 464–470, Springer-Verlag, 2010.
- [3] S. Borgo and O. Kutz, *A General Framework for Shape Similarity*. in *Proc. SHAPES 1.0 - The Shape of Things. Workshop at CONTEXT-11, CEUR-WS*, Vol. 812, 2011.
- [4] F. Cena and S. Likavec and F. Osborne, *Property-based interest propagation in ontology-based user model*. in *Proc. 20th Conference on User Modeling, Adaptation, and Personalization UMAP 2012*, Lecture Notes in Computer Science 7379, pp. 38–50, Springer-Verlag, 2012.
- [5] M. Grüninger and S. Bouafoud, *Thinking Outside (and Inside) the Box*. in *Proc. SHAPES 1.0 - The Shape of Things. Workshop at CONTEXT-11, CEUR-WS*, Vol. 812, 2011.
- [6] J. Hois and M. Bhatt and O. Kutz, *Modular Ontologies for Architectural Design*. in *Proc. 2009 Conference on Formal Ontologies Meet Industry*, pp. 66–77, IOS Press, 2009.
- [7] J. Jiang and D. Conrath, *Semantic similarity based on corpus statistics and lexical taxonomy*. in *Proc. International Conference on Research in Computational Linguistics*, pp. 19–33, 1997.
- [8] P. Jurewicz, *Building Shape Ontology, Organising, Visualising and Presenting Building Shape with Digital Tools*, Diplomarbeit, Technischen Universität Wien, 2005.

<sup>5</sup>PIEMONTE project: People Interaction with Enhanced Multimodal Object for a New Territory Experience, financed by Regione Piemonte, in the context of Converging Technologies.

- [9] M. Niknam and C. Kemke, *Modeling Shapes and Graphics Concepts in an Ontology*. in *Proc. SHAPES 1.0 - The Shape of Things. Workshop at CONTEXT-11*, CEUR-WS, Vol. 812, 2011.
- [10] L. Console and F. Antonelli and G. Biamino and F. Carmagnola and F. Cena and E. Chiabrando and V. Cuciti and M. Demichelis and F. Fassio and F. Franceschi and R. Furnari and C. Gena and M. Geymonat and P. Grimaldi and P. Grillo and E. Guercio and S. Likavec and I. Lombardi and D. Mana and A. Marcengo and M. Mioli and M. Mirabelli and M. Perrero and C. Picardi and F. Protti and A. Rapp and R. Sandon and R. Simeoni and D. Theseider Dupré and I. Torre and A. Toso and F. Torta and F. Vernerio, *WantEat: interacting with social netnetworks of intelligent things and people in the world of gastronomy*, *ACM Transactions on Interactive Intelligent Systems*, Accepted for publication, 2013.
- [11] R. Rada and H. Mili and E. Bicknell and M. Blettner, *Development and application of a metric on semantic nets*, *IEEE Transactions on Systems Management and Cybernetics*, 19(1):1730, 1989.
- [12] P. Resnik, *Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language*, *Journal of Artificial Intelligence Research*, 11:95130, 1998.
- [13] O. Symonova and M. S. Dao and G. Ucelli and R. De Amicis, *Ontology Based Shape Annotation and Retrieval*. in *Proc. ECAI06 International Workshop on Contexts and Ontologies: Theory, Practice and Applications*, 2006.
- [14] A. Tversky, *Features of similarity*, *Psychological Review*, 84(4):327–352, 1977.



# The Shape of Absolute Coincidences. Salmon's Interactive Fork Model as Shape of Coincidental Processes.

Alessandra MELAS  
*University of Sassari, Italy*

**Abstract.** According to a particular view, chance events are not uncaused but they are simply the result of intersecting causal lines. More precisely, the intersections between different processes that belong to independent causal chains are the origin of accidental events, called *absolute coincidences*.

This paper provides a new account devoted to showing the strong relation between absolute coincidences and Salmon's interactive fork criterion, in an attempt to endorse the idea that coincidences can be shaped in terms of a causal model.

**Keywords.** Absolute coincidences, causal, interactive forks, shape

## Introduction

As for the word “chance”, the word “coincidence” is used to indicate many different things. The present study, however, considers only a particular type of coincidences, namely what is known as *absolute coincidences*.

According to Jacques Monod, absolute coincidences are the result of intersections between different processes that belong to totally independent *causal* chains<sup>1</sup>:

Mais dans d'autres situations, la notion de hasard prend une signification essentielle et non plus simplement opérationnelle. C'est le cas, par exemple, de ce que l'on peut appeler les “coincidences absolues”, c'est-à-dire celles qui résultent de l'intersection de deux chaînes causales totalement indépendantes l'une de l'autre.<sup>2</sup>

I will call that type of coincidences “*causal* absolute coincidences”.

The first part of the present work provides a precise definition of *causal* absolute coincidences. The second one presents the strong relation between *causal* absolute coincidences and Salmon's common cause model. As we will see, *causal* absolute coincidences are events that can be divided into intersecting causal components, and

---

<sup>1</sup> In *Metaphysics*, Aristotle already maintains the fact the existence of *per accidens* causes is a sign of the existence of *per se* causes. In commenting on Aristotle's *Metaphysics*, Saint Thomas also says that if we treat accidental beings as things produced by *per se* causes, many things may be by accident, such as the meeting of independent causal lines. Although very important for the philosophical historiography, I leave out from this discussion the Aristotelian-Thomistic notion of chance. For an extended enquiry see [1].

<sup>2</sup> [2], p. 128.

that intersection gives origin to consequences. Equally, Salmon's interactive forks are characterized by two, or more, intersecting processes and two, or more, ensuing processes. It seems that *causal* absolute coincidences can be represented by an *x*-shape, as with Salmon's interactive forks.

The core idea of this paper is that representing and reasoning with the shape of coincidental phenomena is essential to understand those phenomena.

The fact that the DNA molecule has the shape of a double helix is crucial to understand how it functions. The same is true of coincidences: the fact that coincidences have a particular shape is crucial to understand how they work.

### 1. *Causal* Absolute Coincidences: a Definition

As we have already seen, *causal* absolute coincidences are not uncaused, but they are simply the effect of the intersection of independent causal processes.

We can explain the independence between different processes, *A* and *B*, that belong to independent causal chains in the following terms:

- 1) *A* and *B* are independent if they are statistically independent, so that:

$$P(A/\neg B) = P(A/B) = P(A) \quad (1)$$

and

$$P(B/\neg A) = P(B/A) = P(B) \quad (2)$$

- 2) The statistical independence between *A* and *B* is not due to a common cause in their past<sup>3</sup>.

To clarify this point let us consider Monod's example as it is represented in Figure 1. Doctor Dupont is going to visit a patient for the first time. In the meanwhile, Mr Dubois is fixing a roof in the same area. When doctor Dupont comes across Dubois' work site, Dubois' hammer falls inadvertently down and the trajectory of the hammer intersects the trajectory of doctor Dupont, who dies<sup>4</sup>.

In Figure 1:

$$P(A/\neg B) = P(A/B) = P(A) \quad (1)$$

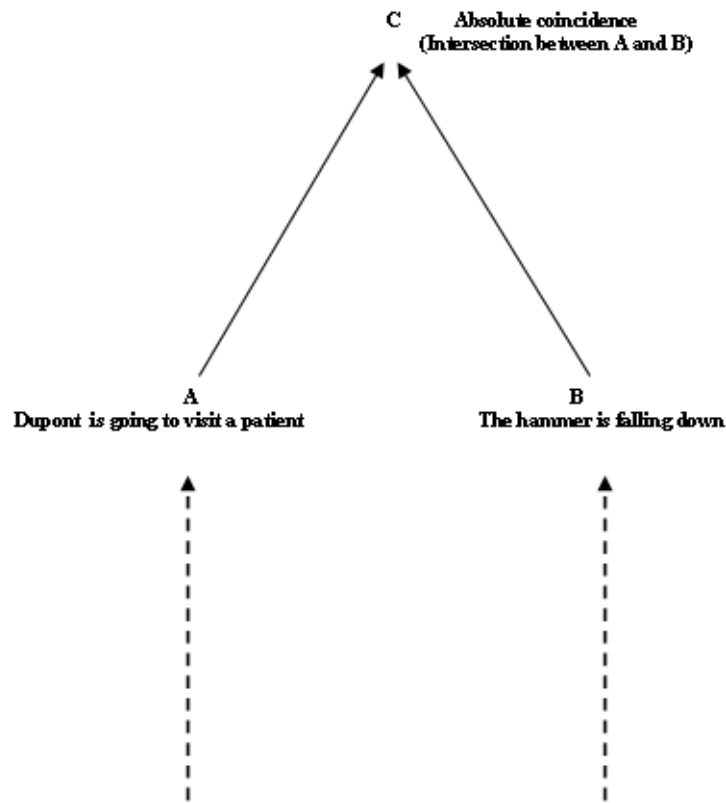
and

$$P(B/\neg A) = P(B/A) = P(B) \quad (2)$$

---

<sup>3</sup> I want to avoid the case in which an event *X* is a screening-off common cause of the two processes *A* and *B*. In that case, according to Reichenbach's screening-off condition, given *X*, *A* and *B* are independent of each other. For a more extended discussion on Reichenbach's screening-off condition see [3].

<sup>4</sup> [2], p. 128.



**Figure 1.** Monod's example of absolute coincidences

That is, the fact that doctor Dupont goes to visit his patient is statistically independent of the fact that the hammer falls down, and the fact that the hammer falls down is statistically independent of the fact that doctor Dupont goes to visit his patient.

The two dotted lines in Figure 1 represent the two independent causal histories of A and B.

To sum up, *causal* absolute coincidences are events that can be divided into components independently produced by some causal factor and those components join together.

Let us consider another example. Suppose I am watching a TV programme on Boris Pasternak. In the meanwhile my best friend, without knowing what I am doing and without knowing anything about that TV programme, is reading doctor Zhivago.

We would say that it is a coincidence that at the same time (but in different places) my friend and I are doing something that concerns Boris Pasternak.

However, since there is no physical direct *interaction* between the two coincidental processes, someone may conclude from this example that something is a coincidence only in the eye of the beholder.

Hence, we need a more precise definition of what a *causal* absolute coincidence is:

*We speak of a causal absolute coincidence whenever there is an intersection, which is also a physical interaction, of two or more statistically independent causal processes in exactly the same space and at exactly the same time.*

A good example could be Monod's one: in that case the two independent causal processes intersect, and physically interact, in exactly the same space and at exactly the same time. No one may conclude that the intersection is a coincidence only in the eye of the beholder<sup>5</sup>.

Moreover, in such cases the intersection gives origin to some consequence, as it is illustrated in Figure 2.

## **2. The Connection Between Salmon's Interactive Fork Criterion and Causal Absolute Coincidences**

*Causal* absolute coincidences are events that can be divided into components which intersect (and also interact) in certain spaces and at certain times and, as we have already seen, that intersection gives origin to consequences.

This section is devoted to showing the strong relation between *causal* absolute coincidences and Salmon's common cause model. More precisely, after an overview of Salmon's model, I will show that coincidences can be entirely described in terms of interactive forks<sup>6</sup>.

### *2.1 Salmon's Interactive Fork Model*

As Salmon says:

[...] Consider a simple example. Two pool balls, the cue ball and the 8-ball, lie upon a pool table. A relative novice attempts a shot that is intended to put the 8-ball into one of the far corner pockets, but given the positions of the balls, if the 8-ball falls into one corner pocket, the cue ball is almost certain to go into the other far corner pocket, resulting in a "scratch". Let A stand for the 8-ball dropping into the one corner pocket, let B stand for the cue ball dropping into the other corner pocket, and let C stand for the collision between the cue ball and the 8-ball that occurs when the player executes the shot. We may reasonably assume that the probability of the 8-ball going into the pocket is also about 1/2 if the player tries the shot, and the probability of the cue ball going into the pocket is also about 1/2. It is immediately evident that A, B, and C do not constitute a conjunctive fork, for C does not screen off A and B from one another. Given that the shot is attempted, the probability that the cue ball will fall into the pocket (approximately 1/2) is not equal to the probability that the cue ball will go into the pocket, given that the shot has been attempted and that the 8-ball has dropped into the other far corner pocket (approximately 1).<sup>7</sup>

---

<sup>5</sup> One may say that coincidences are also unexpected events, so that one could ask whether a non-expected intersection, between processes that belong to independent causal chains, is still a coincidence. I find this point very interesting, but I will leave it out of this discussion.

<sup>6</sup> Although a very similar work can be done using Bayesian networks, I will leave that discussion for another paper.

<sup>7</sup> [4], pp. 168-169.

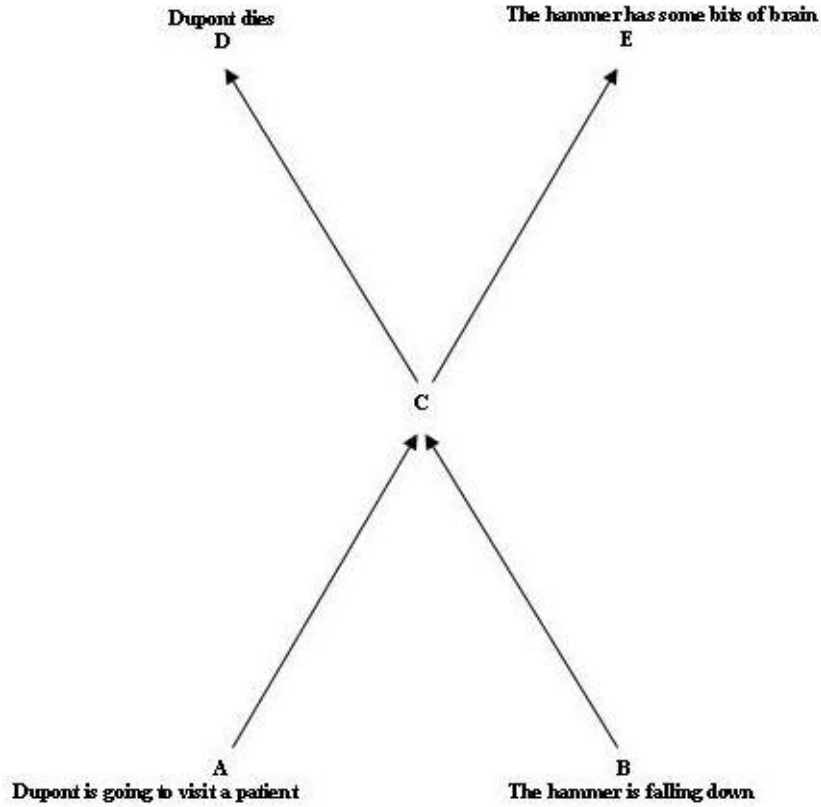


Figure 2. Consequences of coincidences

Salmon's interactive forks are considered as spatio-temporal intersections between two processes. The space-time diagram of the interactive forks has the shape of an  $x$ , as illustrated in Figure 3.

In Figure 3,  $P_1$  and  $P_2$  are two processes which intersect in  $C$ .  $E_1$  and  $E_2$  are the two emerging processes from the intersection.

The intersection produces an interaction and the interaction, according to Reichenbach's mark criterion<sup>8</sup>, has the capacity to produce changes in the properties of the two separate processes that come from the intersection.

In order to characterize Salmon's forks we need the following condition:

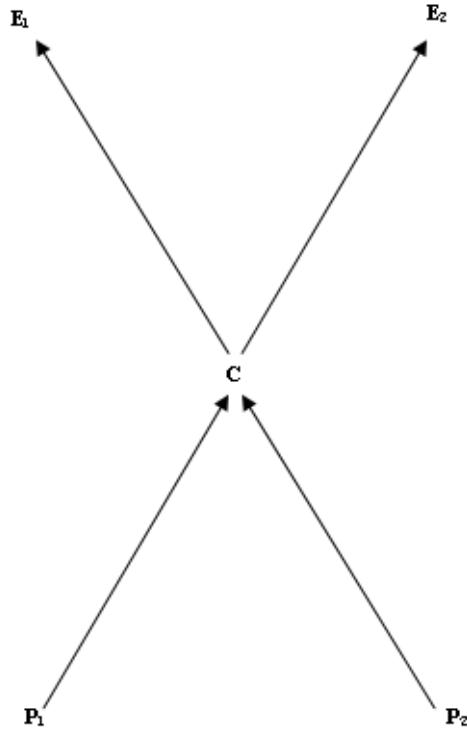
$$P(E_1 \wedge E_2/C) > P(E_1/C) \times P(E_2/C) \quad (3)$$

The intersection in  $C$  makes the two effects statistically dependent on each other. The interactive forks are considered as spatio-temporal intersections that in general violate Reichenbach's screening-off condition<sup>9</sup>.

<sup>8</sup> For a more extended discussion on Reichenbach's mark criterion see [3].

<sup>9</sup> For a more extended discussion on Reichenbach's screening-off condition see [3].





**Figure 3.** Salmon's interactive fork

However, the screening-off condition may represent a limit case for interactive forks. There seems to be a kind of screening-off that is valid for some macroscopic interactive forks<sup>10</sup> and it can be represented by the following condition:

$$P(E_1 \wedge E_2/C) = P(E_1/C) \times P(E_2/C) = 1 \quad (4)$$

which violates the relation (3). Salmon calls this sort of forks "perfect forks". According to Salmon, we can say:

The main point to be made concerning perfect forks is that when the probabilities take on the limiting values, it is impossible to tell from the statistical relationships alone whether the fork should be considered interactive or conjunctive.<sup>11</sup>

Which is, then, the most important difference between conjunctive<sup>12</sup> and interactive forks?

The answer is: the shape of the fork, as we can read from the following quotation:

<sup>10</sup> [4], pp. 177-178.

<sup>11</sup> [4], p. 178.

<sup>12</sup> For a detailed discussion on conjunctive forks see [3].

[...] For it seems essential to have two processes going in and two processes coming out in order to explain the idea of mutual modification.<sup>13</sup>

In recognizing interactive forks, the shape of the phenomena is more relevant than statistical relationships.

Moreover, according to the conjunctive fork model, the common cause can produce one of its effects without producing the other one, and vice versa. According to the interactive fork model, the common cause (the intersection) *cannot* produce one of its effects without producing the other one, and vice versa.

Furthermore, as a result of the physical interaction, in the case of interactive forks, the two emerging processes have something (a mark) of each other and this is not true for the case of conjunctive forks.

## 2.2 The Coincidences *x*-Shape and Interactive Forks

Interactive forks are characterized by two, or more, intersecting (and interacting) processes and two, or more, ensuing processes. Equally, absolute coincidences are events that can be divided into components which intersect (and interact) in certain spaces and at certain times and that intersection gives origin to consequences. It seems that *causal* absolute coincidences can be represented by an *x*-shape, like interactive forks. I will call this *x*-shape *coincidences x-shape*, and I will try to investigate its relation with *x*-shapes which characterize interactive forks.

### 2.2.1 The First Part of the Coincidences *x*-Shape and Interactive Forks

In Salmon we find some instances in which the two intersecting processes,  $P_1$  and  $P_2$ , are statistically independent of each other, as in our definition of *causal* absolute coincidences. We can quote the following example:

In every day life, when we talk about cause-effect relations, we think typically (though not necessarily invariably) of situations in which one event (which we call the cause) is linked to another event (which we call the effect) by means of a causal process. Each of the two events in this relation is an interaction between two (or more) intersecting processes. We say for example, that the window was broken by boys playing baseball. In this situation there is a collision of a bat with the ball (an interactive fork), the motion of the ball through space (a causal process), and a collision of the ball with the window (an interactive fork).<sup>14</sup>

According to this example, the ball travelling towards the window and the window are two intersecting processes; the collision between them represents the intersection. In this case, the two intersecting processes belong to independent causal chains, since the fact that the window is there is statistically independent of the fact that the ball travels towards the window, and vice versa.

Salmon does not care about the distinction between intersecting processes that belong to independent causal chains and intersecting processes that belong to non-independent causal chains. This is not a relevant point for a good definition of interactive forks.

Interactive forks are not a sufficient requirement to say that an interaction happens by coincidence, since a fork can always be found in which the interacting processes do

---

<sup>13</sup> [4], p. 182.

<sup>14</sup> [4], p. 178.

not belong to independent causal chains<sup>15</sup>. However, cases of intersecting processes that belong to independent causal chains are not ruled out by Salmon's account, and Salmon's model can be used to describe such cases.

The first part of the *coincidences x-shape* can be described with Salmon's model.

### 2.2.2 The Second Part of the Coincidences x-Shape and Interactive Forks

Let us consider again Monod's example, as it is represented in Figure 2. In that case we may reasonably assume that the probability of doctor Dupont dying  $D$  is of  $1/2$  if there is a collision between the hammer and doctor Dupont's head  $C$ , and that the probability of the hammer having some bits of brain  $E$  is also of about  $1/2$ . In this case, it is evident that  $D$  and  $E$ , given  $C$ , are statistically dependent on each other:

$$P(D \wedge E/C) > P(D/C) \times P(E/C) \quad (5)$$

Given that the hammer collides with doctor Dupont's head, the probability that the hammer will have some bits of brain (approximately  $1/2$ ) is not equal to the probability that the hammer will have bits of brain, given that the hammer has collided with doctor Dupont's head and doctor Dupont has died [ $P(E/C \wedge D) \approx 1$ ]. That is:

$$P(E/C \wedge D) > P(E/C) \quad (6)$$

Moreover, given the collision between the hammer and doctor Dupont's head, the probability that doctor Dupont will die (approximately  $1/2$ ) is not equal to the probability that the doctor will die, given that the hammer has collided with doctor Dupont's head and the hammer has some bits of brain [ $P(D/C \wedge E) \approx 1$ ]. That is:

$$P(D/C \wedge E) > P(D/C) \quad (7)$$

The intersection in  $C$  makes the two effects,  $D$  and  $E$ , statistically dependent, like in Salmon's example of the two billiard balls.

However, as we have already seen, the condition of screening-off represents a limit case for interactive forks. There seems to be a kind of screening-off which is valid for some macroscopic interactive forks. That limit case is represented by the following condition:

$$P(E_1 \wedge E_2/C) = P(E_1/C) \times P(E_2/C) = 1 \quad (4)$$

Consider the two billiard balls example once more. Suppose that our novice returns to attempt another shot from time to time. Since practice helps improve one's skills to perfection, the novice becomes so good that he can invariably make the cue ball and the 8-ball collide ( $C$ ) in the manner that the 8-ball drops into one of the far corner pockets ( $E_1$ ) and the cue ball goes into the other far corner pocket ( $E_2$ ). We may reasonably assume that the probability of the 8-ball going into the pocket is about 1 if the player tries the shot, and the probability of the cue ball going into the pocket is also

---

<sup>15</sup> The aim of this paper is to show that interactive forks are a *necessary* condition for coincidental phenomena.

about 1. It is immediately evident that *C* does screen *E*<sub>1</sub> and *E*<sub>2</sub> off from one another. Up until the moment when our player has perfected his technique, the results of his shots exemplified interactive forks, and it would be absurd to claim that when he achieves perfection, the collision of the two balls no longer constitutes a causal interaction, but it must now be considered as a conjunctive fork. It is an arithmetical accident that, when perfection occurs, the equation (4) is fulfilled while the inequality (3) must be violated.

Let us go back to Monod's example once again. We may reasonably assume that the probability of doctor Dupont dying (*D*) is of about 1, if there is a collision between the hammer and doctor Dupont's head (*C*), and that the probability of the hammer having some bits of brain (*E*) is also of about 1. In this case, it is evident that *C* screens *D* and *E* off from one another:

$$P(D \wedge E/C) = P(D/C) \times P(E/C) = 1 \quad (8)$$

However, it would be absurd to claim that the collision between the hammer and doctor Dupont's head no longer constitutes a causal interaction, but must now be considered as a conjunctive fork. According to this last example, the common cause (*C*) cannot produce one of its effect without producing also the other one, and vice versa. Moreover, the two emerging processes (*D* and *E*) have something of each other. It is an arithmetical accident that, when perfection occurs, the equation (4) is fulfilled while the inequality (3) is violated.

According to what I said in this section, even the last part of the *coincidences x-shape* can be described using Salmon's interactive fork model.

### 3. Conclusion

To conclude, the primary reasons for saying that interactive forks can describe *causal* absolute coincidences are:

1. Interactive forks are characterized by two, or more, intersecting (and interacting) processes and two, or more, ensuing processes. Equally, absolute coincidences are events that can be divided into components which intersect (and interact) in a given space and at a given time, and that intersection gives origin to consequences. *Causal* absolute coincidences can be represented by an *x-shape*, as with interactive forks.

2. Cases of intersecting processes that belong to independent causal chains are not ruled out by Salmon's account and Salmon's model can be used to describe such cases. The first part of the *coincidences x-shape* can be described by Salmon's criterion.

3. Salmon's interactive fork model can be easily used to describe the last part of the *coincidences x-shape*.

Finally, we can conclude not only there is a connection between *causal* absolute coincidences and the Principle of Causality, according to which whatever begins to exist has a cause, but also the possibility to describe them in terms of some causal model.

## References

- [1] Julienne Junkersfeld, M. (1945) *The Aristotelian-Thomistic Concept of Chance*. University of Notre Dame.
- [2] Monod, J. (1970) *Le Hasard et La Nécessité. Essai sur la philosophie naturelle de la biologie moderne*. Éditions du Seuil.
- [3] Reichenbach, H. (1956) *The Direction of Time*. University of California Press.
- [4] Salmon, W.C. (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.

# The Role of Shape in Problem-Solving Activities in Mathematics

José Gustavo MORALES<sup>a,1</sup> and Matías SARACHO<sup>b,2</sup>

<sup>a</sup>*National University of Cordoba / CONICET*

<sup>b</sup>*National University of Cordoba*

**Abstract.** In our paper we will rely on research by Grosholz (2007) considering her thesis of the irreducibility of iconic representation in mathematics. Against this background, our aim will be to discuss the epistemic value of “shape” or iconicity in diagrammatic representations in geometry. We show that iconic aspects of diagrams reveal structural relations underlying the method to solve quadrature problems developed by Leibniz (1675/76). As a concluding remark, we shall argue that in retrieving the information embedded in a diagram the reader must establish a meaningful relationship between the information supplied by the diagram and the relevant background knowledge which often remains implicit.

**Keywords.** Iconic Representation, Diagrams, Visualization, Leibniz, General Method, Background Knowledge.

## Introduction

In our paper we rely on research by Grosholz (2007) considering, in particular, her thesis of the irreducibility of iconic representation in mathematics. Against this background, our aim is to discuss the epistemic value of “shape” or iconicity in the representations of diagrams in the case of geometry. In order to illustrate our point, we bring in a case-study selected from Leibniz’s work with diagrams in problem-solving activities in connection with a “master problem, the Squaring of the Circle – or the precise determination of the area of the circle”, a problem which remains insoluble by ruler and compass construction within Euclidean geometry.<sup>3</sup> Our main reason to focus on Leibniz is as follows. On the one hand, throughout his work as a mathematician, Leibniz relies on a variety of tools which display rich iconic aspects in the implementation of problem-solving activities. On the other hand, it is precisely in the case of geometry where Leibniz makes important contributions. Reasoning with diagrams plays a central role in this particular case. In order to solve certain geometrical problems which could not be solved within the framework of Euclidean geometry, Leibniz devises a method that proceeds by transforming a certain mathematically intractable curve into a more tractable curve which is amenable to

---

<sup>1</sup> José Gustavo Morales, National University of Cordoba, School of Philosophy, Cordoba (5000), Argentina; E-mail: gust.914@gmail.com.

<sup>2</sup> Matías Saracho, National University of Cordoba, School of Philosophy, Cordoba (5000), Argentina; E-mail: matias.m00@gmail.com

<sup>3</sup> See [3, p. 36].

calculation. This method is sometimes called the method of “transmutation” as it is based upon the transformation of one curvilinear figure into another.

For Leibniz depending upon the context of research some methodological tools are more fruitful than others, moreover, simplicity and economy is also amongst the epistemic virtues guiding the design of methods for problem-solving activities. In our case-study, we show how Leibniz devises a method which allows him to re-conceive a given curve by “transforming” it into a more tractable curve as part of his strategy to calculate the area of curves that may contain irrational numbers (the real number  $\pi$  in the case of the circle). In particular, we aim to show that iconic aspects of diagrams reveal structural relations underlying the process of “transformation” developed by Leibniz in *Quadrature arithmetique du cercle, de la ellipse et de l’ hyperbole* (1675/76).<sup>4</sup>

## 1. The Idea of “Shape” As Iconic Representation

Let us start by focusing on the idea of “shape” in the sense of “iconic representation”. Representations may be iconic, symbolic and indexical depending upon their role in reasoning with signs in specific contexts of work.<sup>5</sup> According to the traditional view representations are iconic when they resemble the things they represent. In many cases this characterization appears as doubtful because of its appeal to a vague idea of similarity which would seem untenable when representations of numbers are involved. But Grosholz argues that in mathematics iconicity is often an irreducible ingredient, as she writes,

In many cases, the iconic representation is indispensable. This is often, though not always, because shape is irreducible; in many important cases, the canonical representation of a mathematical entity is or involves a basic geometrical figure. At the same time, representations that are ‘faithful to’ the things they represent may often be quite symbolic, and the likenesses they manifest may not be inherently visual or spatial, though the representations are, and articulate likeness by visual or spatial means [3, p. 262].

In order to determine whether a representation is iconic or symbolic, the context of research with its fundamental background knowledge needs to be taken into account in each particular case, in other words, iconicity cannot simply be read off the representation in isolation of the context of use. We find here a more subtle understanding of “iconicity” than the traditional view. Let us focus on the idea that representations “articulate likeness by visual or spatial means”. Grosholz suggests that even highly abstract symbolic reasoning goes hand in hand with certain forms of visualizations.

Giardino (2010) offers a useful characterization of the cognitive activity of “visualizing” in the formal sciences. In visualizing, she explains, we are decoding articulated information which is embedded in a representation, such articulation is a

---

<sup>4</sup> In this paper, we shall be refereeing to the French translation (Parmentier 2004) of Leibniz original text *De Quadratura Arithmetica* (1675/76).

<sup>5</sup> The distinction goes back to Charles Peirce’s theory of signs. For a brief discussion of this distinction, see [3, p. 25].

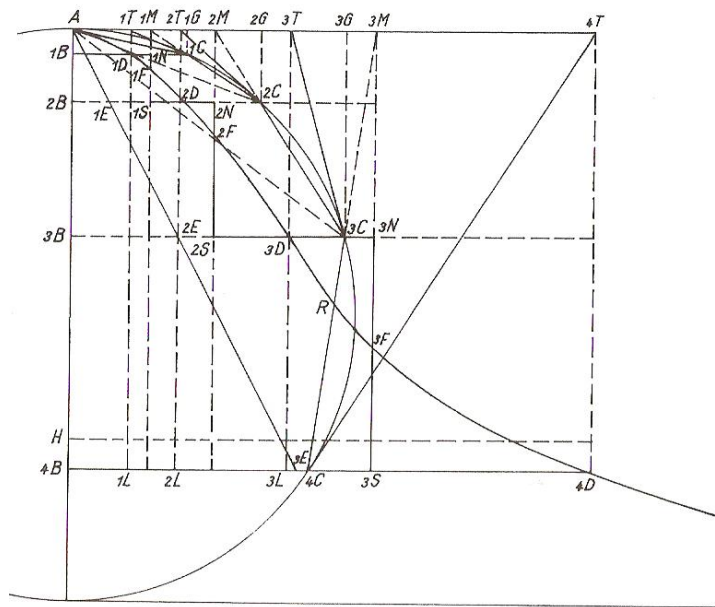
specific kind of spatial organization that lends unicity to a representation turning it intelligible. In other words, spatial organization is not just a matter of physical display on the surface (paper or table) but “intelligible spatiality” which may require substantial background knowledge:

(...) to give a visualization is to give a contribution to the organization of the available information (...) in visualizing, we are referring also to background knowledge with the aim of getting to a global and synoptic representation of the problem [1, p. 37].

According to this perspective, the ability to read off what is referred to in a representation depends on some background knowledge and expertise of the reader. Such cognitive act is successful only if the user is able to decode the encrypted information of a representation while establishing a meaningful relationship between the representation and the relevant background knowledge which often remains implicit. The starting point of this process is brought about by representations that are iconic in a rudimentary way, namely, they have spatial isolation and organize information by spatial and visual means; and they are indivisible things. In the next section we turn to our case study taken from Leibniz’s work in geometry which we hope will help to illustrate some of the above considerations.

## 2. Our Case Study - Leibniz’s *De Quadratura Arithmetica* (1675/76)

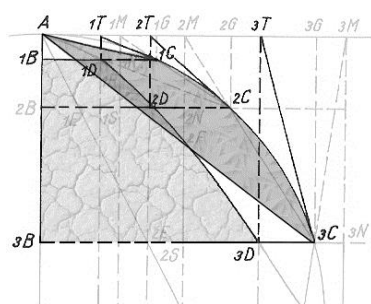
In *Quadrature arithmétique du cercle, de l'ellipse et de l'hyperbole* [8] Leibniz provides a general method whereby “quadrature” problems for curvilinear figures can be solved. The first seven propositions of this work form a unity and as Leibniz himself emphasizes, Proposition 7 is the “fruit” of all that has gone before [8, p. 35]. In this context of work, Leibniz presents the reader a diagram (Fig. 1).



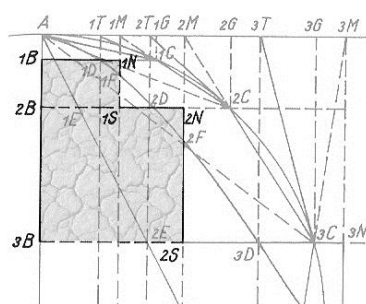


**Figure 1.** Leibniz *Quadrature arithmetique du cercle, de l'ellipse et de l'hyperbole* 1675/76, [8, p. 65].

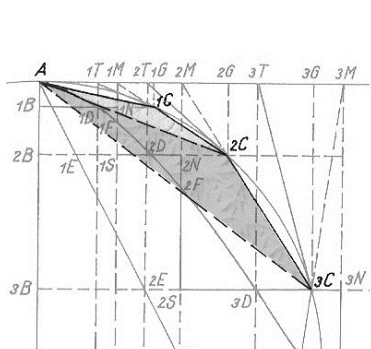
While for the untrained eye this diagram appears as a set of highly entangled shapes, for Leibniz, the diagram should offer the reader an overall assessment of the way his proposed method works. In order to show the most salient aspects of Leibniz's method we shall try to make more explicit some of the features displayed in Figure 1. We proceed to put the original diagram "under the microscope" dissecting it into four diagrams (Figures 2-5). This will allow us to see some of the most relevant steps involved in the resolution of the problem under consideration. These visualizations together with the indications as to how to "read" Figures 2-5 may then be seen as offering a brief outline of Leibniz's method.



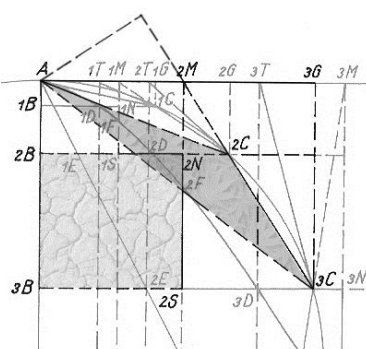
**Figure 2**



**Figure 3**



**Figure 4**



**Figure 5**

Leibniz aims to show that the area of a curvilinear figure C – which cannot be calculated - may be determined by constructing a second figure D, whose area can be calculated. A crucial step in Leibniz reasoning relies upon certain geometrical results known since Euclid which allow us to assume that the ratio between C and D is known to us. This step in the reasoning is represented in the diagram by two different shapes that we have highlighted in Figure 2. On the one hand, we see an enclosed area

delimited by segments  $A_1C$ ,  $A_3C$  and the arc  ${}_1C_2C_3C$  – which represents the area  $C$ , unknown to us. On the other hand, we see another enclosed area delimited by segments  ${}_1B_1D$ ,  ${}_1B_3B$ ,  ${}_3B_3D$  and the curve  ${}_1D_2D_3D$  which represents the area of the second figure  $D$ . Finally, we can also see some specific lines that represent geometrical relations between both figures according to Euclidean geometry.<sup>6</sup>

With a view to determine the area of curvilinear figure  $C$  we first need to find the area of figure  $D$ . Leibniz proceeds to decompose  $D$  into a finite number of elemental parts - the rectangles  ${}_1N_1B_2B_1S$  and  ${}_2N_2B_3B_2S$  - which are then added up. We have highlighted this procedure in Figure 3. As we can also see the sum of rectangles makes up a new shape or figure which Leibniz calls “*espace gradiforme*”.<sup>7</sup> At this stage of the reasoning, the construction of such “space” is crucial for Leibniz’s problem-solving strategy. Instead of an exact calculation of the area of  $D$ , Leibniz approximates the area of  $D$  by calculating the area of such “*espace gradiforme*”, so that the difference between both figures will be less than any assignable number.

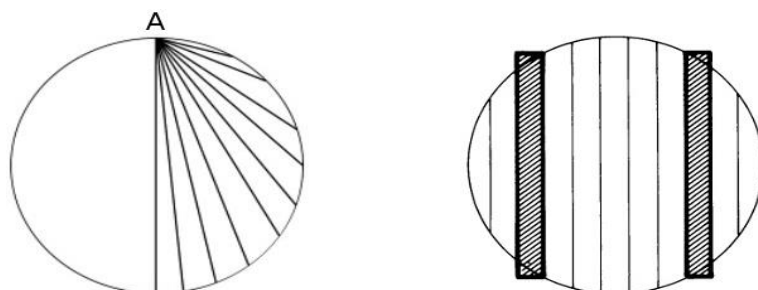
Next, the newly constructed “*espace gradiforme*” is transposed upon figure  $C$  (See Figures 4 and 5). This procedure can be described in two steps.

The first step consists in decomposing the curvilinear figure  $C$  into “triangles” which we highlighted in Figure 4. Note that the number of triangles will be greater than any arbitrarily assignable number as it is possible to decompose the figure into arbitrarily many triangles where the whole set of triangles has the single vertex  $A$ . Here Leibniz takes distance from other techniques used at the time. While Cavalieri, for instance, often decomposed curvilinear figures into parallelograms, Leibniz proceeds to resolve the problem by decomposing curvilinear figures into triangles (for an illustration of this difference see Figure 6). Accordingly, instead of rectangles or parallelograms, the elemental parts in this case will be triangles, as Leibniz points out in Scholium 1 of the treatise:

(...) on peut en effet également décomposer en triangles des figures curvilignes qu’à l’exemple d’autres grands savants Cavalieri ne décomposait souvent qu’en parallélogrammes, sans utiliser, à ma connaissance, une résolution générale en triangles [8, p. 39].

<sup>6</sup> Leibniz relies upon a generalization of Euclid’s Elements (Proposition 1, Book I) to justify his reasoning when assuming that the “triangle”  $A_2C_3C$  equals one half of the “rectangle”  ${}_2B_2N.{}_2B_3B$  ( $A_2C_3C = 1/2 {}_2B_2N.{}_2B_3B$ ). See [9].

<sup>7</sup> The expression Leibniz used in the original Latin is “*spatium gradiforme*” [8, p. 69].



**Figure 6.** Leibniz's method as opposed to Cavalieri's method.

The second step consists in the construction of the “*espace gradiforme*” upon C (See Figure 5). To this end, Leibniz uses the rectangle with sides  ${}_2B_2N$  and  ${}_2N_2S$  which can be constructed from a given triangle  $A_2C_3C$  relying on certain well-established geometrical relations which hold so that the ratio between the areas of figures C and D is  $\frac{1}{2}$ . It is precisely in this context where Leibniz relies upon results already established by Euclid.<sup>8</sup>

Let us now return to Leibniz's original diagram corresponding to Proposition 7 (See Figure 1). With Euclid's results concerning structural relations between two types of shapes - triangles and rectangles – in mind, we are justified to establish a correlation between triangles  $A_1C_2C$ ,  $A_2C_3C$ ,... and corresponding rectangles  ${}_1B_1N_2B_1S$ ,  ${}_2B_2N_3B_2S$ ,.... For instance, the triangle  $A_2C_3C$  corresponds to the rectangle  ${}_2D_2B_3B_2S$ . Next, we recall that the area of figure D can be approximated by the sum of the (finite number of) elemental parts – rectangles – the original figure D was decomposed into.

Finally, the area of the curvilinear figure C can be calculated by applying the ratio of  $\frac{1}{2}$  upon the area of figure D. According to Leibniz, the calculation obtained by this method is not exact but one may consider it is a precise determination of the area of the curvilinear figure C. To sum up, it is by recognizing certain geometrical relations holding between triangles and rectangles that one can see that the precise determination of the area of the curvilinear shape will depend upon the value of the approximation of the area of D. The latter, in turn, can be calculated on the basis of the “*espace gradiforme*”, the new shape designed by Leibniz which is required to approximate the value of D.

### 3. Concluding Remarks

In this section we finally consider some of the requirements which are imposed upon the reader in order to be able to perform the relevant “cognitive act” of successfully decoding a visualization that includes “shapes” in the context of problem solving activities in mathematics. Again, we shall focus on the diagram of our case-study (Figure 1).

Diagrams are shapes that represent by spatial and visual means. Their intelligibility partly depends on their integrity and shape, features that make a diagram intrinsically iconic. But diagrams also often combine, as Grosholz argues, iconic aspects with

<sup>8</sup> See footnote 6 (above).

symbolic ingredients. If diagrams were just iconic, they would be but a copy – a more or less faithful picture - of what we intent to refer to. However, diagrams are inherently general, a drawing of, say, curvilinear shape without being just a drawing of a particular curve on this particular page of a text. On the one hand, diagrams resemble a particular shape, on the other hand, they represent a whole set of (instances of) a certain shape and are in this sense general. To clarify this feature of diagrams we distinguish following M. Giaquinto between “discrete” and “indiscrete” representations,

(...) diagrams very frequently do represent their objects as having properties that, though not ruled out by the specification, are not demanded by it. In fact this is often unavoidable. Verbal descriptions can be discrete, in that they supply no more information than is needed. But visual representations are typically indiscrete, because for many properties or kinds  $F$ , a visual representation cannot represent something as being  $F$  without representing it as being  $F$  in a particular way [1, p. 28].

“Indiscrete representations” as opposed to “discrete representations” are representations that represent by spatial and visual means including the combination of iconic aspects as well as symbolic ingredients. As a consequence of this important feature of diagrams, it follows that both particular instances and generality go hand in hand. Returning to our case-study and Leibniz’s diagram, we may offer the following three observations in this regard:

- The diagram that goes with proposition 7 (Figure 1) exhibits a circular shape. We may consider that Leibniz’s method to calculate the area for this curvilinear shape works only for this particular curve. But Leibniz intends to use his method as a general method so as to include any curvilinear shape, as he writes in the Scholium to proposition XI:

La proposition 7 m’a fourni le moyen de construire une infinité de figures de longueur infinie égales au double d’un segment ou d’un secteur (...) d’une courbe donnée quelconque, et ceci d’une infinité de manières (Leibniz 1676, p. 97).<sup>9</sup>

- In the diagram (Figure 1) the curvilinear shape  $C$  is actually divided into only four points, namely,  ${}_1C_2C_3C_4C$ . However, it is possible to divide the arc  $C$  into as many points as we want.
- If the number of points is large enough, the diagram will be less faithful to the particular instance that it pretends to represent and when the magnitude of segment  $A_1C$  is less than any assignable number, we have the limit-case. At this point, the space  ${}_1CA_3C_2C$  (called “triligne” by Leibniz) can be assumed as a space composed by curve  ${}_1C_2C_3C$  and the straight line  $A_3C$  (called “secteur” by Leibniz)<sup>10</sup>.

<sup>9</sup> Leibniz specifies the class of curves which fall under the domain of application of his method in Proposition 6 of *Quadrature arithmetique*.

<sup>10</sup> See [8, p. 97].

Note that in our case-study, the reader has to select only part of the information furnished by the diagram; he/she has to be able to discern the relevant information contained in the diagram in the light of the problem under consideration. In particular, it is necessary to distinguish in the diagram between iconic ingredients and symbolic ingredients. What exactly is required of the reader to be able to decode the relevant information encrypted in the diagram? To answer this question we return here to Giardino's observation that "to give a visualization is to give a contribution to the organization of the available information". First the reader needs to consider the context in which the diagram is inserted. As already noted, part of the context is made explicit by remarks written in natural language as it is the case in the written text accompanying the diagram [8, pp. 65, 67]. In the written text, Leibniz explains how to construct the diagram he shows together with Proposition 7 (our Figure 1). But such description is hardly enough, as the reader still needs to rely on substantial information – background knowledge concerning relevant chapters of the history of geometry – in order to get "a global and synoptic representation of the problem". However the relevant background knowledge cannot be made fully explicit, at least not "all at once". The expertise of the community of mathematicians which includes different traditions of research and, in a broad sense, the history of mathematics, provides different tools and techniques which need to be acquired by teaching and learning. For instance, in our case-study Leibniz's diagram relies heavily on procedures and techniques whose origin goes back to Euclid and Archimedes but also recalls the work of some of his contemporaries such as Cavalieri's "theorem of indivisibles" and Pascal's "characteristic triangle", which is used by Leibniz in order to "transform" triangles into rectangles. Finally, as we may divide the arc C into "as many points as we want", sometimes the diagram is meant to be read as an infinitesimal configuration and at this point the symbolic dimension of the diagram comes into play so that in each case the trained eye of the reader will be required to be able to recognize the roles of these different dimensions.

## References

- [1] Giaquinto, M. (2008) Visualizing in Mathematics, in P. Mancosu, *The Philosophy of Mathematical Practice*, Oxford University Press, 2008
- [2] Giardino, V. (2010), Intuition and Visualization in Mathematical Problem Solving, Dordrecht, Springer, *Topoi* (2010) 29:29–39.
- [3] Grosholz E. (2007) *Representation and productive ambiguity in mathematics and the sciences*. Oxford New York, 2007.
- [4] Grosholz, E. (2010) Reference and Analysis: The Representation of Time in Galileo, Newton, and Leibniz, *Arthur O. Lovejoy Lecture*, Journal of the History of Ideas, Volume 72, Number 3 (July 2011), pp. 333-350.
- [5] Ippoliti, E. (2008) *Inferenze ampliative. Visualizzazione, analogia e rappresentazioni multiple*, Morrisville, North Carolina (USA), 2008.
- [6] Lakoff, G., Núñez (2000), *Where Mathematics Comes From. How the Embodied Mind Brings Mathematics into Being*. Basic Books, The Perseus Books Group, 2000.
- [7] Leibniz, W. G., (1849 - 1863), *Mathematische Schriften*, vol. I – VII, ed. Gerhardt Hildesheim: Olms, 1962.
- [8] Leibniz, G.W. (1675/1676), *Quadrature arithmétique du cercle, de l'ellipse et de l'hyperbole*, Paris, *Librairie Philosophique J. VRIN*, French Translation by Marc Parmentier, 2004.
- [9] Parmentier, M. (2001), Démonstrations et infiniment petits dans la Quadratura arithmetica de Leibniz, *Rev. Hist. Sci.* 54/3. 275-289.

# Shaping up: The Phenotypic Quality Ontology and Cross Sections

Robert J. ROVETTO<sup>1</sup>

**Abstract.** The Phenotypic Quality Ontology (PATO) uses the notion of a cross section to relate two- and three-dimensional shapes and to describe the shape of biological entities. What is a cross-section? What is a truthful ontological account of cross sections? In this communication I explore potential answers to these questions, approaching the task from philosophical and ontological perspectives, and provide a preliminary examination of the PATO shape hierarchy. Some critical observations, questions, and suggestions for the shape portion of PATO are presented.

**Keywords.** Shape, ontology, applied ontology, philosophical ontology, biomedical ontology, artifacts, cross section, philosophy of mathematics, Phenotypic Quality Ontology, PATO, ontology of shape

## Introduction

A preliminary examination of the shape hierarchy in the Phenotypic Quality Ontology (PATO) [1] was conducted. Herein I present some critical observations, questions and suggestions for the ontology. The developers of PATO employ the notion of a cross section in order to describe the shapes of biological entities. What is a cross section, and what is an ontological description of them? Potential answers to these questions are explored in the discussion on the ontology of cross sections.

During my analysis of the shape taxonomy, I organized all PATO shape classes in a spreadsheet document, using temporary identifiers for each, and devised a simple comment-code system. Sibling classes were grouped according to the level they are subsumed under the root shape class, ‘shape’, with the groupings being called the subsumption or nesting level. That is, all siblings classes subsumed once by ‘shape’ (subsumption level 1) I labeled ‘Level 1 terms’ and so forth for all the sibling terms subsumed twice (or once under Level 1 terms).

In what follows, key terms and phrases are italicized. Potential ontological categories are also emphasized in boldface, and PATO classes are enclosed in single quotes. Occasionally the latter are emphasized and prefixed with the name of the respective ontology as in: PATO:**shape**. “Category”, “class” and “universal” are considered interchangeable, each essentially reflecting general(izable) entities. Similarly for “particular”, “instance” and “individual”. In section 1 I introduce PATO, the `has_cross_section` relation, and present some critical observations thereof. In section 2 I provide an ontological discussion of cross sections.

---

<sup>1</sup> Corresponding Author. University at Buffalo, The State University of New York, Department of Philosophy (2011 MA graduate), United States; Email: rrovetto@buffalo.edu.

## 1. PATO: The Phenotypic Quality Ontology

PATO is an applied ontology<sup>2</sup> primarily intended to describe the phenotypic qualities (traits, more specifically) of organisms. Qualities are most commonly and broadly understood in their normal sense as properties, characteristics or attributes of things. For PATO, qualities are more formally understood in the sense of BFO:**quality** of the Basic Formal Ontology in [2] (version 1.1) or [3] (version 2.0). As such, they are a type of dependent entity<sup>3</sup>. Qualities can also be formally described in the sense of DOLCE:**quality** of the Descriptive Ontology for Linguistic and Cognitive Engineering [4], which interestingly differentiates qualities from properties by making them particulars and universals, respectively.

PATO is designed to be used with ontologies of quality-bearing entities [5]. By explicitly mentioning this (type of entity) PATO can be seen as taking a particular stance on what and how things exist, that is, a metaphysical stance on the world. Since the ontology commonly uses [2] as its top-level ontology<sup>4</sup>, one can argue that the former essentially adopts the metaphysical assumptions and theory of the latter. If true, this is an important point to remember for any implicit or explicit, contested or controversial, ontological assertions in the top-level. It is significant for persons and theories that hold the ontology of the world to be different. This point would appear to apply to any ontology using an upper-level with a thoroughly worked-out metaphysical, ontological or philosophical theory.

PATO is included in the Open Biological and Biomedical Ontologies Foundry (OBO Foundry) [6]. It is intended to be used with other biomedical ontologies, such as the Gene Ontology [7], and is used for phenotype annotation. According to [8], the Neuroscience Information Framework [9] and the Influenza Ontology [10], among others, use PATO. Examples of PATO classes include ‘shape’, ‘size’, ‘texture’, ‘structure’, ‘physical object quality’, ‘cellular quality’, ‘functionality’, ‘process quality’, and ‘intensity’ to name a few. PATO encompasses a broad range of so-called qualities from the highly general to the more domain-specific, and is ambitious in that respect.<sup>5</sup> Figure 1 displays a sample of the shape taxonomy in PATO as it was structured at the time of this examination. Indentation indicates subsumption (the *is\_a* relation). The first two shape classes, for example, can be read as “aliform is a type of 2-D shape”.

---

<sup>2</sup> An applied ontology is an ontology used in or for a field of inquiry or domain science, such as biology or physics. It has a computer-readable format to foster, among other things, automated reasoning, and semantic annotation and interoperability. A (scientific) ontology is ideally a philosophically well-founded theory of the kinds of entities and relationships in the world. Domain ontologies reflect an aspect or portion of the world that is the focus of study for a domain science. For more see [2], [4] and [22].

<sup>3</sup> A dependent entity is an entity that depends on others for its existence. Qualities/properties, being commonly-held examples, are described as being existentially dependent on their material bearer.

<sup>4</sup> A top-level ontology is one whose ontological categories are of the highest generality possible to accurately describe the world or some field of inquiry. Each lower-level category of a lower-level or domain-ontology that uses a top- or upper-level has some upper-level category subsuming it.

<sup>5</sup> With classes such as ‘physical object quality’ and ‘process quality’, both highly general and applicable to different types of entities, the subject matter of PATO is clearly broader than the domain of biology. The abiotic is represented in the ontology as well.

```

2-D shape
  aliform
3-D shape
  convex 3-D shape
  concave 3-D shape
angular
arrow-shaped
auriculate
bent
branchiness

```

**Figure 1.** A portion of PATO’s shape taxonomy.

Shapes are generally held to be examples of qualities of things. This is intuitively accepted when considering the shape of material objects, but not so obvious when considering the abstract or ideal shapes of geometry. The shapes of material objects (*physical/material shapes*) do not typically exhibit the flawless symmetry of *geometric shapes*. The distinction between the two, discussed in [11], may help (I) specialize a shape category for applied ontologies such as PATO, and (II) form a correct ontology of shape. PATO shape qualities, at least in the sense of [2, 3], are essentially equivalent to the former: the shape of material entities in the world.

### 1.1. Some Initial Observations

An initial observation of PATO’s shape hierarchy is that of a largely flat one that could benefit from more structuring and category assessment. This includes determining whether the categories are representative of shapes (qualities). For example, two superclasses of PATO:shape are ‘Morphology’ and ‘physical object quality’, a classification that is not entirely correct. Strictly speaking, morphology is a field or subject of study, not the shape quality of a physical object. The intended reading and meaning is likely “Morphological quality”. If so, then an appropriate change in the class name or definition to reflect this would better afford semantic transparency. That is, grasping the meaning of the class from its name will be more apparent.

Although there are similar concerns with other PATO classes and definitions, including those that do not appear to denote shapes (PATO:robust for example), I will focus on observations and implications related PATO:has\_cross\_section. Before discussing this relation, I will briefly present two potential strategies for organizing the shape categories.

There are a number of classes whose names end with “-shaped” or “-like”, as in PATO:hourglass-shaped, ‘snowman-shaped’, ‘spoon-shaped’ and ‘brush-like shape’. Some of them can undoubtedly be removed, renamed or substituted. If two things are accomplished—a distinction between artifactual and non-artifactual shapes, and an unproblematic general account of artifactuality—then these classes can arguably be grouped accordingly under (Non-)Artifactual shape classes.

PATO does not have a class for spheres, and perhaps justly so. Spheres are geometric shapes and seemingly outside the scope of an ontology whose primary domain is that of phenotypic qualities of biological entities. A *Sphere* class would not fit because (a) geometric shapes are purportedly abstract in nature, (b) their instances are not exhibited by material objects, and (c) the shape qualities of PATO require a material bearer.

Spheres are mentioned, however, in the definitions of PATO:spheroid, the parent class (the immediate superclass) of PATO:spherical. Now if it were possible for some



material entity to exhibit a perfectly spherical shape—giving credence to the idea of perfect material instances of geometric shape universals—then these classes are general enough to encompass those shape qualities in addition to approximate spheres or sphere-like shapes, such as the oblate spheroidal shape quality of the Earth.

Proceeding with the understanding that any given physical object does not, in fact, have a perfect geometric shape quality, we have the following. Assuming the distinction between geometric and physical shapes is both ontologically correct and viable for an applied ontology, differentiating between *Sphere* and *Spherical* categories would reflect the distinction and may prove useful. Instances of a *Sphere* class have the precise geometric properties of a geometric sphere while allowing for variation in scale. For PATO, instances are physical shapes, shape qualities approximating a geometric sphere. A spherical biological cell can be formally related to the geometric sphere via some two-place approximation or resemblance relation.

In noun-based adjectives<sup>6</sup> [12], such as “spheroidal” and “rectangular”, the suffixes “-oidal” (or “-oid”), “-ical”, and “-ar” indicate resemblance (and thus deviation from the ideal). The definition of PATO:**spherical** is in line with, stating “...the bearer’s resembling a ball”<sup>7</sup> (italics added). In terms of class names, those with the above suffixes can be defined as indicating this similarity or approximation. They mark imperfect physical object shape qualities, rather than ideal geometric shapes. This approach would require geometric shape categories in a separate ontology, however.

### 1.2. The PATO cross section relation

To better structure the shape hierarchy, the developers of PATO utilize the notion of a cross section to relate three- and two-dimensional (3-D and 2-D for short) shapes [13]. More specifically, the `has_cross_section` relation, a binary predicate, was introduced to relate 2-D and 3-D shape classes. A description of the 3-D shape of a biological cell, for example, would have a 2-D counterpart marking the cross section of that 3-D cell shape. This ontologically commits PATO to cross sections, yet there is no class or corresponding definiton. Below is the definition of PATO:**has\_cross\_section** with an arbitrary label to the left. I have emphasized the likely unique entity- and category-referring words.

(HCS-Def)      **s3** `has_cross_section` s3 if and only if : there exists some **2d plane** that intersects **the bearer of s3**, and the **impression** of s3 upon that plane has **shape quality s2**.

Example:      a spherical object has the quality of being spherical, and the spherical quality `has_cross_section` round [13]

According to both the definition and the example, the domain and range are 3-D and 2-D shape qualities, respectively, each a subtype of top-level class BFO:**quality**. Table 1 displays the form and a sample subject-predicate-object expression (the row in grey). Based on the phrasing of the relation the range is cross sections, more

<sup>6</sup> Personal communication with Dr. John Corcoran, University at Buffalo, State University of New York.

<sup>7</sup> The use of the word “ball” is ambiguous, but seems to connote material objects, e.g., a soccer ball, rather than a geometric shape. The definition says that a ball *is a* sphere, but it is the soccer ball that resembles a sphere, not vice versa.

specifically. The implication is that there is some equivalency between cross sections and 2-D shapes. For PATO, it would appear that 2-D shapes are cross-sections. The definition of PATO:**2-D shape** classifies cross-sections as 2-D entities, more generally. For both geometric and physical shapes we can ask, Q1: Are cross sections 2-D shapes (or vice versa)?

**Table 1.** The PATO has\_cross\_section relation with domain, range, and working example (grey row)

Domain	Binary Relation	Range
3-D shape quality	has_cross_section	2-D shape quality
Spherical	has_cross_section	Round

Observe that translating HCS-Def completely into first-order predicate logic may require an intersection relation, as well as some relation reflecting the usage of “upon”. In the next subsection I present three critical observations of HCS-Def, while attempting to address them. Each is labeled with the prefix “HCS-O” and a number.

### 1.3. The has\_cross\_section definition – Observations and Questions

(HCS-O1) Assuming there is no typographic error, the definition may be better served if the second instance of “s3”, located on the left side of the bi-conditional, is substituted with “s2”. “S2” is also found at the end of the definition, and likely stands for “2-D shape”. Using the corresponding full PATO class name in the definition is arguably preferable. There are at least two reasons for the substitution, reiterating a point made in subsection 1.1.

The first is that since the range is 2-D shape qualities, the substitution will more easily reflect this, making for a more readable and transparent definition to ontology editors and users. Similar minor changes will make formal definitions more consistent with any informal natural language description, and perhaps encourage the promotion of standards and best practices for applied ontology development and curation.

The second reason is that if both instances of “s3” signify distinct entities of the same type—either 3-D or 2-D shapes—then the relation cannot apply to geometric shapes, nor to the bearer of the shape, mind-external 3-D material objects. A sphere does not have a 3-D cross section, nor does, say, a particular orange. A cross section is commonly, if not always, of a lower-dimensionality than the cross-sectioned entity.

(HCS-O2) 2-D planes are mentioned in HCS-Def, signaling an ontological commitment to planes. In what sense do 2-D planes exist and what kind of entity are they? Do they interact with the bearer of the shape quality, as the definition suggests, and if so, how? The two are clearly of different types.

The 2-D plane is a geometric or mathematical entity, entities not explicitly represented in PATO’s top-level, [2], and thus beyond the scope of PATO. Depending on the nature of geometric/mathematical entities, planes may be a type abstract entity, more generally, or a cognitive entity. [4], and much philosophical literature, supports geometric shapes and mathematical entities as abstract. By contrast, and according to PATO, the bearer of a shape quality is a BFO:**material entity**.

Assuming planes are mathematical entities, how exactly do they interact with material objects? With these questions we find ourselves in the sphere of the philosophy of mathematics in which a central concern is the ontological status of mathematical entities. Are they mind-internal constructions, mind-external spatio-temporal entities, abstract (non-spatio-temporal and non-mental), or otherwise? Is a

*mathematical entity* class to be introduced to subsume a **2-D plane** class? If so, then in what ontology are they to be a part (a mathematical entity ontology, perhaps)? How would this ontology be used with PATO and its top-level?

Consider the following approach. If planes interact with 3-D material bearers, or if cross sections are only created by an intersection (a form of interaction) of distinct entities, then introduce an *intersection* relation, a binary predicate with 2-D planes and 3-D material objects (or shapes qualities) as domain and range. The problem is that we would be hard-pressed to explain how they interact in this way. We do not literally find 2-D planes in the mind-external world, except perhaps in a representational sense.

Within a computational or mathematical system, a 3-D model of a real-world object can be intersected (in whatever equivalent sense) with a 2-D plane. This would be a *computational representation* of an otherwise abstract geometric entity. It makes more sense to ontologically describe the interaction of these entities in a cognitive, mathematical, or computational setting or context.

I am therefore inclined to suggest that only geometric shapes can be intersected in this sense. That is, a *geometric intersection* relation would relate geometric entities. Physical shapes are only subject to intersection with a plane when they are abstracted from the mind-external 3-D material bearer, idealized or represented in a cognitive, mathematical, or computational form. Perhaps the physical shape, so abstracted, can be modeled as a geometric shape.

(HCS-O3) Although we intuitively understand the meaning of “impression” from HCS-Def, it is not described and marks an ontological commitment to impressions. Is an *impression* class, and corresponding logical definition, to be added to PATO? Is it a primitive, the meaning of which comes from our natural-language comprehension of its usage in context? Is the impression identical to the cross-section? Let us examine this and the intersection process implied by HSC-Def a bit further.

The intersection of a 2-D plane with a 3-D shape quality results in a 2-D impression of the latter on the former. So how does the impression relate to the 2-D plane? It is not clear whether the impression is on the plane as a book can be on a table, or whether the impression is a quality of the plane. Both appear implausible.

To say either the shape or the cross section is on the plane is to suggest that the cross section is independent of both the plane and sphere, just as a table and a book are independent, but interacting, entities. Ontologically, there is no independent cross section standing in an *on* relation (in this sense) to the plane. The cross section appears to form by the intersection.

A mereological theory can be used to assert that the sphere and plane partially overlap, and define the cross-section as the region of overlap. This would support the idea that the cross section is not an independently existing entity. Provided a theory of location, part of the 2-D plane can also be described as being co-located with part of the 3-D shape. These formal accounts would more accurately apply to cross sections in a geometric context as types of immaterial, abstract or mathematic entities.

If impressions have 2-D shape qualities, then impressions are quality-bearers. Are we to gather that the 3-D material object is the bearer of the 3-D shape quality, the plane a bearer of the impression, and finally the impression a bearer of the 2-D shape quality? This may be problematic.

According to PATO, shape qualities, not the bearer, have cross sections. Yet according to [2], PATO’s adopted top-level ontology, qualities do not themselves have qualities. At best, any asserted high-order quality is a quality of the bearer. Given this, and assuming cross sections are 2-D shape qualities, the material bearer would have the

cross-section as a quality, not the 3-D shape quality of the bearer as PATO holds (see 1.2 and Table 1). Unless planes and impressions play some role in avoiding this apparent dilemma, a contradiction exists. This is a potential, if not actual, conflict between the ontological theory of the top-level and the lower-level ontology.

From the preceding, we observe that PATO is ontologically committed to cross sections, shape qualities and their bearers, and arguably 2-D planes and impressions upon the planes. We can partially address HCS-O3 by being ontologically parsimonious and avoid commitment to impressions. Rephrasing the definition to avoid use of “impression”, or (perhaps) explicitly stating that the word be understood in an informal manner are possible solutions. HCS-O3 introduces the question, Q2: Do/can both shapes and material objects have cross sections? It appears so.

The next section discusses the ontology of cross sections. I identify some of their general properties, labeling them with the prefix “CS-P” and a number as they appear. Potential definitions/descriptions are marked with “CS-D”. I also present interpretations of cross sections as types of (a) mind-external artifacts, (b) mind-internal entities, (c) mind-external portions of objects, and (d) spatio-temporal regions.

## 2. The Ontology of Cross Sections

What are cross sections, and what is an accurate ontological description of them? We generally conceive of cross sections as either two-dimensional slices/sections of things, or geometric description of things. The former are purportedly parts or types of material entities, and the latter types of representational artifacts in the sense of [14, 15]. In short, cross-sections are 2-D sections or descriptions of 3-D entities. We also understand the notion of cross-sections by an action or process: cutting/slicing or intersecting an entity to expose some internal structure, surface or shape.

Generally speaking, we can section something mind-internally by imagining the cross section of some entity; mind-externally, by performing actual dissections of biological specimens; or virtually/computationally by producing computational representations and simulations. These are three modes of cross-sectioning, the last two of which involve the first. Relevant definitions from both [16] and [17] reflect the idea that a cross section can be an actual cut/cutting of something, or a representation of it.

In medicine, 3-D sections of biological entities—cells, tissue, organs, organisms—are obtained by actual dissections or other cutting/sectioning processes. By contrast, 2-D slices are virtually obtained or simulated with technology such as Computed Axial Tomography in which X-rays scan a biological subject. This scanning process, in effect, produces 2-D slices in a computational or graphical form in order to visually convey information about the interior of the subject.

Numerous slices taken from varying angles and positions are combined to form a 3-D mosaic: a 3-D computational representation of, say, the human brain. A particular brain is virtually sliced or sectioned based on, or by the gathering of, information that is (i) received and subsequently processed by the scanner and associated software, and (ii) transmitted via X-rays. Ultimately, 2- and 3-D images and simulations representing the subject are produced for the medical practitioner to examine. Thus, medical imaging provides us with an example of (representational) cross sections.

Other fields, such as architecture and engineering, use cross sections (or variations thereof) with the aid of software applications such as AutoCad [18] and Photoshop [19]. These applications permit similar constructions and comparable simulated dimensional

(de)compositions to communicate internal and structural information. It is evident that these cross sections have a descriptive power that is highly useful for industry and education. Consider blueprints, floor plans, scientific illustrations, computer simulations, and modeling and manufacturing techniques. Whether in medicine or engineering, from computational cross-sectioning processes—types of artifactual processes—representational artifacts are created.

**(CS-P1)** A cross section is always a cross section of something, just as a physical/material shape as opposed to a geometric shape, is always a shape of the object [11]. Cross sections are always cross sections of some entity. They reflect some aspect of what is sectioned. An individual cross section of a particular brain depends at least on that brain, the persons engaging in the sectioning, and the technology involved. The notion of a cross section presupposes that of something to be sectioned (**CS-P2**).

We consider, cognize, imagine and seek the cross section of an object for a purpose. We consider it to satisfy a particular function. Functions, like qualities, are often ontologically classified as a type of dependent entity. *Medical Cross Sections*, a potential domain-level subcategory of a *Cross Section* category arguably have the function of communicating structural or internal information about the biological subject. These considerations support the idea that (**CS-P3**) cross sections are a type of dependent entity. It remains to be seen precisely what type of dependent entity.

When we imagine the cross section of some real or fictional object we are abstracting from the object. We selectively focus on certain parts, properties or aspects of the object, while ignoring others. This is common to any abstraction process. The mind-internal abstraction yields to our mind's eye a figure—the cross section—whose boundaries are that of the representation of the mind-external object. With no mind-external manifestation of the abstraction—artifacts such as images of the brain—the cross section will remain a mental construction. We can ask Q3: Are cross sections types of artifacts, or do these artifacts represent cross sections?

If we take the position that an image from a medical scan represents a cross section of the brain, rather than being a *computational cross section* of the brain, then we ontologically commit to two types of entities. One is the representation and the other is the represented. The images surely represent some physical part, some collection of properties, some snapshot of an ongoing biological process, or some aspect of the brain, more generally. That aspect is of the internal structure (and goings-on) of the scanned portion of the brain (1) viewed two-dimensionally or in a 2-D form, and (2) relative to some axis (of rotation, of symmetry) or some pre-defined coordinate system.

In short, the cross section could be described as **CS-D1**: a 2-D representation/description of the internal structure (and contents) of the subject at a moment of time and relative to a given axis or coordinate system. If cross sections are created by persons to serve the function of conveying this information, then it is obvious that there is an artifactual element to their nature. Similar to representational artifacts, they could be declared as types of *information artifacts*. Accordingly one may wish to categorize cross sections as *information content entities* in the sense of [20]. Whatever the case, a particular cross section is a dependent entity that conveys and reflects some information, some reality, about what is sectioned.

To cross-section or to section something is to divide or cut it into one or more N-dimensional slices/sections. For 3-D material objects, any actual slices are always of the same dimensionality. There are no 2-D material slices of a brain, for example. Therefore, unless understood as 2-D descriptions or representations of (portions of) things, and unless a cross section category is specialized, it may not be appropriate for

material entities. For geometric shapes, by contrast, the slices are indeed of a lower dimensionality.

The sectioned entity is either actually of a higher dimensionality, or is conceived in the imagination as an entity having a higher dimensionality (**CS-P4**). Implicit in the notion of a cross section is the distinction between (entities of) differing dimensionality (**CS-P5**). For material objects this entails material extension along those dimensions. **CS-P4** implies **CS-P6**: Cross sections are of a lower dimensionality than that which is cross-sectioned. Considerations of dimensions provides an interpretation of cross sections as a 3-D entity viewed from the perspective of a 2-D entity.

### *2.1. Geometry and Cross sections*

There is a distinct geometrical nature to cross-sections. One reason for the dimensional difference between the cross-section(ed) is that a 2-D geometric entity, such as a plane, is used to produce or exhibit the cross section. From the perspective of a plane, 3-D geometric shapes and 3-D material objects are viewed two-dimensionally. They become or are seen as 2-D. Since material objects are not actually cut into 2-D slices, whatever slice is produced is more along the lines of **CS-D1**, a 2-D view, representation or description of the object. Whether the sectioned entity is geometric or material, there is always a conscious involvement or focus (**CS-P7**).

In geometry, the cross section of a 3-D geometric shape (or figure more generally) is the shape produced by the intersection of the 3-D shape with a 2-D plane. A cross section of a sphere is a circle, for example. From the vantage point of a plane, a sphere appears as a circle. For more complex or irregular figures the cross section varies significantly depending on its position and angle through the figure. The position and angle are identifiable by specifying an axis (or coordinate system) through or relative to the geometric object.

Imagine a 2-D plane intersecting a representation of an aircraft. A plane perpendicular to an axis that runs from nose-to-tail will produce a different cross section if it were perpendicular to an axis running from wing-tip to wing-tip. Each will therefore communicate and represent different properties and information about the aircraft. By specifying a rotational axis or coordinate system for the sectioned entity, the orientation of the 2-D plane and resultant varying cross sections can be precisely identified. With this in mind, perhaps **PATO:has\_cross\_section** is not simply a binary predicate but ternary such that: X has\_cross\_section Y at/relative to Z, where Z is some axis, coordinate system or set of coordinates.

The cone and the conic sections are examples of this *cross-sectional variation* for geometric shapes. A plane intersecting a cone at different positions and angles produces different curves: the parabola, the hyperbola and the ellipse (a closed curve), of which the circle is one type. Cross sections would appear to have the same dimensionality as the plane or figure used to geometrically intersect the sectioned shape (**CS-P8**).

One exception to this idea involves planes that are tangential to higher dimensional shapes. Consider the effect on **PATO:has\_cross\_section**. Assuming cross sections of geometric shapes are themselves shapes, and assuming the point (a 0-D entity) is asserted as a type of shape, the following holds. When a plane tangentially touches or intersects any shape—at the apex of a cone or at the pole of a sphere, for example—the cross section at that region is a point. The intersection yields a shape of lower dimensions than the plane. Therefore, if points are accepted as proper cross sections,

then the range of the `has_cross_section` relation must change to cover not only 2-D shapes, but shapes of lower dimensions as well.

Thus far we have observed that from the perspective of a 2-D plane 3-D shapes often vary depending on the orientation of the plane through the entity. In other words, the cross section depends on certain relationships obtaining between specific entities. With this we now explore cross sections and processes further.

## 2.2. Cross-sectioning Processes

As mentioned in section 1, we can describe cross sections in a processual manner. That is, we can formulate an account of intersecting figures as a process. The *geometric intersection process* of a plane with a sphere yields a circular cross section. In an ontology that has *state* as a category, the state of being intersected results from that process. This particular cross section exists when the sphere and the plane stand in one or more specific relations to each other.<sup>8</sup> It exists when shapes are participating or engaging in certain processes manifesting those relations. Although it appears timelessly true that the cross section of a sphere is a circle<sup>9</sup>, one may distinguish between particular and universal cross sections, holding that the former only (come to) exist during an intersection process.

If so, then given the discussion in preceding sections the individual cross section is neither a property of the individual sphere nor of the plane. These (**CS-P9**) instances of geometric cross sections can be ontologically described as *relationally-dependent entities*, perhaps more specifically as *relationally-dependent shapes*. ‘Relational quality’ from [3] is a class that reflects this idea in certain respects, but diverges from the conception by concerning material, not geometric, objects.

Some plane participates in a geometric intersection process with some sphere, the intersection of which is a circle. We call this the cross section. The individual circle is not an independent entity in this context. It is exhibited by the intersection, reflecting some information or properties about the sphere. The information may include the radius, circumference, chord length or area (the cross-sectional area) of the sphere at, or relative to, certain reference points or axes. In this way, we can understand geometric cross sections as intersections in a sense (**CS-P10**).

Accepting CS-P10, we describe geometric cross sections as, **CS-D2A**: the geometric intersection of some N-dimensional shape, in an N-dimensional space, with a 2-D plane. If one wishes to allow for intersections involving figures other than planes, such as the intersection of a sphere with a hyperbola or a cube, we can more generally hold **CS-D2B**: Geometric cross sections are the intersection, or the figure exhibited by the intersection, of some N-dimensional shape with another. What if we wish to analyze the intersection in terms of points or lines?

Descending to a lower geometric granular level, that which is intersected more discretely than the unities that are the sphere and the plane, is that which they have in common: points or lines. They can be described as sharing or having overlapping points when they intersect. Along these lines, geometric cross sections are **CS-D3**: the set of points shared by a 3-D shape and a 2-D plane when the two intersect. Connecting those points forms a continuous figure that is the cross section in question.

<sup>8</sup> Some may wish to call this amalgam a geometric context.

<sup>9</sup> Is this equivalent to saying that the cross section, itself, has a circular quality? Is it a property of spheres that they have circular cross sections?

Finally, observe that a plane continuously moving through a sphere and perpendicular to one of the spheres axes results in a geometrical cross section changing through time. The circle continuously changes its size, diameter and area while essentially preserving its identity as a (relationally-dependent) circle. Imagine the same continuous motion for the aircraft example. At each temporal moment, or at each point along the axis, one can identify either a changing cross section or distinct cross sections. We can therefore describe a cross section at a single moment of time and diachronically.

### 2.3. Cross sections as Spatio-Temporal Entities

Some applied ontologies, such as [2], include two-dimensional spatial regions as an ontological category. With [2] as it's top-level, one can easily hold that PATO implicitly accepts these types of entities as well. Using a *spatial region* category we can form an alternative characterization of a cross section. A cross section is the 2-D portion of space-time, or the 2-D spatial region, whose boundaries are defined by the spatial or material extent of some part of the 3-D material object.

In short, we may be tempted to assert **CS-D4**: cross sections are 2-D portions of space occupied by the sectioned entity. Similarly one may equate cross sections with 2-D material portions/sections of things (**CS-D5**). These would imply that any portion of 2-D space or of the 3-D material object is a cross section. Both CS-D4 and CS-D5 are most likely incorrect characterizations of cross sections.

Cross sections are not literally 2-D material slices of things because, again, 2-D material slices of 3-D material objects are never created. Matter being 3-D, there is always some depth to a material slice. If cross sections are 2-D slices of 3-D material objects, and space(-time) is at least 3- or 4-D, then 2-D material slices or spatial regions will not be found in the ontology of the world: they do not (and cannot?) exist.<sup>10</sup>

At best, an actual *physical/material cross-section* (as with a dissected brain) could be construed as the *surface* of the section. In this case, if surfaces of material objects are included in an ontology as types of 2-D entities, then perhaps these 2-D material surfaces can be equated to 2-D material cross sections. One may propose **CS-D6**: A cross section of a material object is the surface of an internal section (part) of the object, a surface that is orthogonal to a predefined axis or coordinate system, and that is exposed by some sectioning process. This conception is in line with [21], which defines a cross section as the surface or shape that is or would be exposed by cutting something orthogonal to an axis.

We do not observe cross sections of things outside of, and unrelated to human cognition or activity, and thus outside of some artifactual context. They appear to be, in large part, our creations, but creations intentionally reflecting some aspects of other entities. They are often born as mental artifacts that are mind-externally concretized using information about, or acquired from, the entities that are cross-sectioned.

### 3. Closing Remarks

The Phenotypic Quality Ontology uses shape quality categories and the notion of a cross section to describe the shape of biological entities. Drawing from a preliminary

---

<sup>10</sup> If the universe is at least 3-D, then it may be true that all 2-D entities—2-D shapes, 2-D spatial regions, etc.—are all dependent entities, perhaps mental constructions or mathematical tools.



analysis of the shape hierarchy of PATO, I first presented some critical observations, questions and suggestions, focusing on the `has_cross_section` relation. Second, I discussed the ontology of cross sections, identifying general properties, and presenting potential definitions and interpretations of cross sections.

Finer distinctions and greater clarity is needed, specifically with regard to (1) the ontology of geometric/mathematical entities, how they relate to artifacts and material objects, and (2) the relationship between the universal-particular and geometric-physical shape distinctions. With these in place a more accurate ontology of shape may better structure and organize existing applied ontologies that use shape categories.

A point worth reiterating concerning theoretical/philosophical and applied ontology development in general is the comparison of the top- and lower-level ontological theories. For any domain-ontology using a philosophically rigorous top-level it is beneficial for the developers of the former to assess whether they agree with the ontological/metaphysical account of the latter, or whether theirs is consistent with it.

Cross sections have a descriptive power in that they convey information about an entity of a greater dimensionality. They simultaneously reflect properties of the sectioned entity, and have artifactual and relationally-dependent aspects.

## References

- [1] PATO: Phenotypic Quality Ontology. [http://obofoundry.org/wiki/index.php/PATO:Main\\_Page](http://obofoundry.org/wiki/index.php/PATO:Main_Page), last accessed Febuary 2013.
- [2] Basic Formal Ontology. <http://www.ifomis.org/bfo>, last accessed November 2012.
- [3] BFO 2 Reference Document. <http://code.google.com/p/bfo/>, last accessed November 2012.
- [4] DOLCE: Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) <http://www.loa.istc.cnr.it/DOLCE.html>, last accessed November 2012.
- [5] PATO: Phenotypic Quality Ontology. <http://obofoundry.org/wiki/index.php/PATO>About>, last accessed November 2012.
- [6] The Open Biological and Biomedical Ontologies Foundry. <http://www.obofoundry.org/>, last accessed November 2012.
- [7] The Gene Ontology. <http://www.geneontology.org/>, last accessed November 2012.
- [8] The National Center for Biomedical Ontology (NCBO) Biportal <http://bioportal.bioontology.org/ontologies/40960>, last accessed November 2012.
- [9] Neuroscience Information Framework. <http://neuinfo.org/>, last accessed November 2012.
- [10] Influenza Ontology [http://influenzaontologywiki.igs.umaryland.edu/wiki/index.php/Main\\_Page](http://influenzaontologywiki.igs.umaryland.edu/wiki/index.php/Main_Page), last accessed November 2012.
- [11] Rovetto, R. J. The Shape of Shapes: An Ontological Exploration. KR-MED. CEUR, 222 (2011), 57-65.
- [12] Givón, T. Notes on the Semantic Structure of English Adjectives. *Language*, 46 (1970): 4, 816-837.
- [13] PATO. [http://obofoundry.org/wiki/index.php/PATO:Revised\\_2008](http://obofoundry.org/wiki/index.php/PATO:Revised_2008), last accessed November 2012
- [14] Smith, B., Kusnierczyk, W., Schober, D., Ceusters, W. Towards a Reference Terminology for Ontology Research and Development in the Biomedical Domain. KR-MED. CEUR 222 (2006), 57-65.
- [15] Johansson, I., Lynøe, N. *Medicine & Philosophy*. Ontos Verlag Frankfurt, 2008.
- [16] Cross-section. Cambridge Dictionary Online: Free English Dictionary and Thesaurus. Cambridge University Press. <http://dictionary.cambridge.org/dictionary/>, last accessed November 2012.
- [17] Cross section. Merriam Webster's Collegiate Dictionary Online. <http://www.merriam-webster.com/dictionary/cross%20section>, last accessed November 2012.
- [18] AutoCAD. <http://usa.autodesk.com/autocad/>, last accessed November 2012
- [19] Adobe Photoshop. <http://www.adobe.com/products/photoshop.html>, last accessed November 2012
- [20] Information Artifact Ontology. <http://code.google.com/p/information-artifact-ontology/>, last accessed November 2012.
- [21] Cross section. Oxford Dictionaries Online. Oxford University Press. [http://oxforddictionaries.com/us/definition/american\\_english/cross%2Bsection](http://oxforddictionaries.com/us/definition/american_english/cross%2Bsection), last accessed January 2013.
- [22] Munn, K., Smith, B.: *Applied Ontology: An Introduction*. Ontos Verlag Frankfurt, 2008.

# Dynamic Assembly of Figures in Visuospatial Reasoning

Stefan SCHNEIDER <sup>a,1</sup>, Ulf KRUMNACK <sup>a</sup> and Kai-Uwe KÜHNBERGER <sup>a</sup>

<sup>a</sup>*Institute of Cognitive Science, University of Osnabrück*

**Abstract.** An exploratory, qualitative experiment sheds light on the depictive theory of mental imagery. The study analyzes the very operations subjects undertake when solving visuospatial tasks. Preliminary results indicate that subjects do not make use of stable mental images: instead, they continuously assemble and re-assemble different perspectives through the guidance of heuristics and prototypes. These observations allow a reinterpretation of mental imagery. We want to forward the hypotheses that a) the assembly process itself is of much higher importance than usually acknowledged; b) that an assembled perspective (or figure) is defined by one's orientation towards certain operations; and c), that heuristics and prototypes are instantiated by a heterarchical organization of mental operations.

**Keywords.** mental imagery, visual synthesis, model construction, spatial orientation

## Introduction

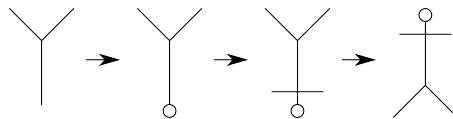
What characteristics does the cognitive representation of a shape have? How do humans reason *about* and *with* shapes? Besides from subjective reports of the appearance and the usage of mental images, the existence of a faculty of imagery – however it may be conceptualized – has been proven to be functionally involved in cognitive processes. Evidence is for example presented from studies showing that humans can mentally synthesize and subsequently recognize figures. In the paradigmatic experiment of Finke et al. [1], subjects were asked to compose objects following verbal instructions like the following:

Imagine the letter Y. Put a small circle at the bottom of it. Add a horizontal line halfway up. Now rotate the figure 180 degrees. ([1, p. 62])

Two opposed conceptualizations of mental imagery are established: the depiction theory and the description theory. These have been forwarded by Kosslyn [2] and Pylyshyn [3] respectively, who strongly opposed the other's account; the dispute is known as the *imagery debate*. Despite substantial doubts whether the depictive approach is computationally feasible and cognitively plausible [3], the intuitive apprehension that in *imagery* we operate with *images* of some sort, has become a hardened paradigm and thereby obstructs alternative conceptualizations. In the following we will present such an

---

<sup>1</sup>Corresponding Author: Stefan Schneider, Institute of Cognitive Science, University of Osnabrück, Germany. E-mail: stefschn@uni-osnabrueck.de.



**Figure 1.** A figure that can be easily mentally synthesized and recognized. Re-drawn from Finke et al. [1].

attempt. First, we introduce the depictive approach, and then confront it with hypotheses derived from a reinvestigation of the very operations subjects execute when solving visuospatial tasks.

### 1. Stable images at your disposal?

In the experiments of [1], the idea to test for possible reinterpretation is cleverly chosen: One cannot argue that recognition of the figure rests upon semantic information in the given description (e.g. as in “Imagine two legs on a torso, add two arms and a head: what have you got?”), but has to rely on cognitive operations with visuospatial information. Figure 1 illustrates the intended assembly of a stick-figure. The reader probably has made his/her own interpretation, possibly similar, by now. It is important to note that subjects in the experiment only received the verbal instruction and did not use any external aids. The visual presence of the depiction in Figure 1 might therefore be misleading, and this, so we want to suggest, shaped the folk-psychological intuition about mental images as *pictures in the head*, and laid the foundation of the dominant depictive theory of imagery.

According to Tversky “[t]here are two basic tenets of the approach, one regarding representations and the other regarding operations on representations: that mental images resemble percepts, and that mental transformations on images resemble observable changes in things in the world, as in mental rotation, or perceptual processes performed on things in the world, as in mental scanning” [4, p. 211]. In a nutshell, the depictive account says that in solving visuospatial problems, one can distinguish two types of operations: (a) image-construction operations, and (b) subsequent inspection of the image, analogous to visual perception.

So far the theory. There is something in disorder, though. Apart from transformations of the image, Kosslyn highlights that images *fade* – “they are transient and begin to decay as soon as they are activated” [2, p. 50]. A frequent *reconstruction* is necessary, which eventually brings about transformations of the imagined objects or a new perspective. The fading of images stands in opposition to their asserted analogue nature to external images (or scenes). This seems to be a flaw in what otherwise could be a pretty useful imagery faculty. We suggest that shifting the research focus away from a supposed presence of images, and towards the process of *reconstructing* these will be of great value.

### 2. Operations in imagery. Current research

In an exploratory protocol study, six subjects had to solve visuospatial tasks like the above with a slightly higher level of complexity. Interviews lasted for about one hour, and were guided according to an interview protocol devised by Petitmengin [5]. In the

analysis, reports about the fleetingness of the images became overly present (which subjects often experienced as a nuisance), and accordingly the necessity to continuously assemble and re-assemble the intended figure. This becomes evident if the reader considers the following task:

Imagine a square. Imagine its vertical lines being separated each into three parts of equal length. Now connect the resulting division marks by two horizontal lines, such that the square is divided into three identical rectangles. Now add the two diagonals to the square. How many triangles does the pattern contain? (Adapted from Wiener [6, p. 82f].)

Even though the elements of the figure are easy to imagine, and the whole figure is not more complicated than the stick-figure, it is very difficult to find all the triangles in this task (there are much more than one initially might think). Interestingly, constructing the figure seemed feasible. A typical report of the resulting image reads like this:

*Well ... I have to ... juggle a bit to keep the components in focus. Everything is there. But I successively imagine how the horizontal lines join the square, and how the diagonals cross these lines, that's all a bit waggly.*

But when asked to count the triangles, *the picture changes*. A typical report is:

*I lose the whole image. It isn't as clear as before any more. (Can you still examine it?) Yes, but I have to build it up anew. And in considering a detail of the image I ... somehow have to retrace the constituents, for example putting in the diagonal again.*

This change of the subject's representation of the figure suggests that the task – in this case, constructing the image versus finding triangles – guides the process. In the latter case, for example, subjects rely on the given elements of the figure and assemble subsets of these in order to construct triangles. That is, instead of *looking at the image* and *recognizing* triangles, subjects actively try to construct triangles from the given material. The following transcript illustrates such an attempt; the subject has the idea that there could be another type of triangle and actively tries to assemble it:

*There is another idea - to search for another line with which the diagonal can form a triangle - there have to be more, because the diagonals cross the lines of the two horizontal lines. I begin in the upper left corner and ... trace the diagonal ... and it goes a level deeper, and forms a triangle where it crosses the second horizontal line. I didn't really search for the horizontal line, but looked how the diagonal runs, whether it could form a triangle somewhere. ... I never have the whole image, but only - well, I often lose the whole image. I trace lines, of which I should already know how it runs. ... Then I felt reminded of the symmetry, and the thought comes up that there have to be four of them.*

The subject has an idea – we want to call it a prototype – of a certain type of triangle which only half of the subjects found. It is being formed by one of the vertical sides, an adjacent diagonal, and the horizontal line which is further away from it; this description would already suffice to deduce a triangle. But for the subject, this is only a hypotheses so far which he tries to instantiate: he traces the diagonal and tries to coordinate it with the horizontal line. But he cannot simply see this line and the intersection. Instead, he makes the horizontal line up again from the initial figure, and coordinates the intersection with the diagonal. Another interesting aspect about his description is the intuition of symmetry, which provides a heuristic to look out for more triangles.

We want to briefly paraphrase our findings. Focusing on the assemble operations, we can see that different perspectives are being assembled over and over. This process



**Figure 2.** Another figure that can be easily synthesized. The task though consists in (mentally!) recognizing *how many triangles the figure contains*. This should be easy! Is it?. From Wiener [6].

is guided in a very fine-grained manner by a scaffolding of the task structure (triangles) and on-the-fly produced prototypes and heuristics. We consider these to be ontologically equivalent, because all guide the assembly of elements of the initial figure and thereby provide a tightly meshed, heterarchical organization of mental operations. Because all of these guides path the way to assemble components, we suggest that an assembled perspective (or figure) is defined by one's orientation towards certain operations. Eventually, an assembled perspective confirms the instantiation of a certain property or relation.

### 3. Summary and hypotheses

Based on our exploration of visuospatial reasoning we want to propose a renewed interpretation of mental imagery. In imagery, one actively attempts to assemble an intended structure; but *one does not simply see it*. One can indeed think of the whole process as being constituted through operations of type (a) and (b) from above, but with a reversed order. We do not make up an image to see something, but we want to see something, and in the absence of a visual stimulus we have to construct an orientation that fits.

Is the visual metaphor of depictive theory – imagery as internalized perception – therefore invalid? This only holds if we think of seeing as a one-way, passive uptake of input. But in an account of active perception, like that of Rensink, he states that one looks at scenes in a structural manner: “scene representations are no [...] structures built up from eye movements and attentional shifts, but rather, are structures that guide such activities [7, p. 36]”.

With regard to the top-down guidance that we found in imagery we are sympathetic to such a conceptualization of vision. But do we therefore have to rephrase and say, somewhat awkward, perception is externalized imagery? Based on our observations there would be a missing piece, which are the overly present dynamic processes of assembling figures. We think it holds that imagery has its ontogenetic foundation in intentionally guided visual operations, and go with Piaget [8] who envisages that based on ones representation of actions, one eventually develops the capability to substitute external through internal processes. The dynamic assembly of figures might be just the organism's way around the lack of available sensor stimulation.

Our results are preliminary. In order to further our understanding of assembly processes, we are currently executing a study on the development of spatial orientation in familiar environments. Our approach also shows promising relations to a recent study, where [9] investigated assembly processes in the apparently unrelated faculty of mathematic cognition.

## References

- [1] Ronald A. Finke, Steven Pinker, and Marthaj Farah. Reinterpreting Visual Patterns in Mental Imagery. *Cognitive Science*, 13:51–78, 1989.
- [2] Stephen Michael Kosslyn. The Medium and the Message in Mental Imagery: A Theory. *Psychological Review*, 88(1):46–66, 1981.
- [3] Zenon W. Pylyshyn. What the Mind’s Eye Tells the Mind’s Brain. *Psychological Bulletin*, 80(1), 1973.
- [4] Barbara Tversky. Visuospatial Reasoning. In Keith J. Holyoak and Robert G. Morrison, editors, *The Cambridge handbook of thinking and reasoning*, chapter 10, pages 209–240. Cambridge University Press, 2005.
- [5] Claire Petitmengin. Describing one’s subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, 5:229–269, 2006.
- [6] Oswald Wiener. Materialien zu meinem Buch Vorstellungen, 2000.
- [7] Ronald A Rensink. The dynamic representation of scenes. *Visual Cognition*, 7(1-3):17–42, 2000.
- [8] Jean Piaget. *The Psychology of Intelligence*. The International Library of Psychology, Philosophy and Scientific Method. Littlefield, Adams & Co., Paterson, New Jersey, 1960.
- [9] Stefan Schneider. *The cognitive structure of the enumeration algorithm: Making up a positional number system*. Master’s thesis, University of Osnabrück, 2012.



# Declarative Computing with Shapes and Shadows

Carl SCHULTZ and Mehul BHATT  
Cognitive Systems (CoSy)  
University of Bremen, Germany

## Abstract.

We present a preliminary concept and a prototypical implementation of a declarative computing framework that is capable of reasoning about 3D physical entities, and the shadows that they cast in open or uniformly lit environments. For this paper, we restrict our scope of ‘uniform lighting’ to sunlight, and its incidence on a given geospatially and temporally referenced location.

The model extends traditional techniques from computational geometry and computer graphics that are primarily motivated by simulation or visualisation. In particular, our declarative framework is capable of deriving and reasoning about the objects and their cast shadows in a knowledge processing sense, e.g., involving qualitative abstraction and semantic specification of requirements, query capability, ensuring conceptual consistency of design requirements. Our ontology of objects and shadows, and the resulting computational framework serves as a foundational engine for high-level conceptual (spatial) design assistance technology.

The capabilities demonstrated in this paper are aimed at applications in spatial design, chiefly encompassing Computer-Aided Architecture Design (CAAD), Urban Planning, and Interior Design.

**Keywords.** declarative languages, knowledge representation and reasoning, geometric and spatial representation and reasoning, computational geometry, shadows, CAAD, design

## 1. Introduction

The way in which direct sunlight falls on surfaces in the built environment has a tremendous impact on its atmosphere, character, and affordances. Consider the role of sunlight in fostering a golden autumnal scene, the mood established by acutely angled sun rays in the early morning, or headache-inducing glare in a work place.

The *absence* of direct sunlight is shadow. Shadows and sunlight partition *empty space*; they are not objects in the sense of having a material extension in the way that walls, doors, and other physical objects do. Yet architects are centrally concerned with the play between shadows and sunlight, and reason about the physical *geometric forms* and arrangements of shadows and sunlight, for example, to achieve a visual balance, to focus



or accentuate aspects of the design, or to create a visual flow through a hierarchy of illumination.

Questions surrounding the behaviour of sunlight come to the fore in urban-scale designs. Architects need to determine how the orientation, shape, and positioning of buildings and large environment features influence the geometric forms of shadows. How can we manipulate the objects in the environment to achieve a desired atmosphere through shadows and sunlight? Would it be possible to manipulate shadows *directly* in our design?

Various numerical techniques have been developed for computing the effects of lighting in an environment. One method is detailed *ray tracing* where a large number of simulation rays are emitted from the light source and reflected from intercepting surfaces.<sup>1</sup> While being very accurate and precise, there are significant limitations with these detailed numerical approaches.

Firstly, they require an enormous amount of computational resources, and may take hours or even days to produce a lighting model for large-scale urban designs. This makes it infeasible to repeatedly re-run the simulation after making minor changes to the design.

Secondly, the results do not emphasise the essential form of shadows *as objects*. For example, they may produce a complex point cloud of luminance values that requires yet further complex calculations to answer basic questions about sunlight and shadow. By applying a uniformly high degree of numerical precision to every aspect of the building design, these methods hide those critical *qualitative* aspects of the design that play the most important role in natural lighting.

Thirdly, these methods require a detailed numerical building design before the lighting model can be generated. But very often, particularly in the early stages of a design, many numerical details are simply not available, such as the exact lengths of certain walls or the precise orientation of certain buildings.

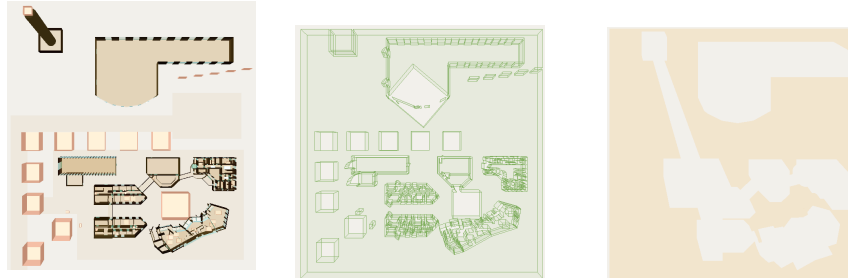
We present a qualitative approach for generating shadows and sunlight regions as first-class objects using appropriate abstractions of building information models and natural lighting. Our objective is to extend and enrich standing design information models with different types of natural lighting *space*. We accomplish this within the paradigm of constraint logic programming so as to support declarative queries and high-level analysis about sunlight and shadows.

## 2. The Shape of Shadows

Seen as mathematical objects, shadows are a product of the interaction between opaque geometric forms and light. The placing and orienting of walls and windows is *sculpting* the forms of shadows and sunlight, and thus shaping yet another layer of experience from the environment, beyond the material objects and the perceived empty space [Bhatt et al., 2012]. Figure 1 comparatively illustrates these three layers of a scene.

---

<sup>1</sup>Each ray might have a maximum number of allowable reflections before the simulation of the ray is terminated. Properties of the ray are recorded, such as the complete path or just the intersection points, and this data is used to generate a detailed surface lighting model.



**Figure 1.** A design represented as the material, *perceived empty space*, and resulting *shadows*.

### 2.1. Spatial Patterns and the Subjective Experience

Within each layer the designer can specify spatial patterns that they deem to evoke salient user experiences or induce certain behaviours. Consider the following brief extracts from Seidler’s opening statements about the role of sunlight and shadow in architecture [Seidler, 1959]:

*“Solid form is accentuated and added to by the shadow it casts which recalls design form.”*

*“Spidery elements, adjuncts to buildings and sculpture increase the interest of their own forms by the complex pattern of their shadows which must be considered as an integral part of the elements that cast them.”*

*“The oscillation between light and shade gives richness to building in contrast [..]”*

In these excerpts the architect identifies a number of relationships: shadows *accentuate* the experience of material forms; complex patterns of shadow *increase interest* in architectural forms; oscillation patterns add *richness* to the experience of the environment.

For example, a small patch of sunlight can be accentuated by a much larger region of surrounding shadow. Such a pattern may function as a visual attractor for user wayfinding, and the visual emphasis and focus may provide a heightened sense of *excitement* and *drama*. We argue that these patterns are *qualitative*, rather than metric, in nature, and that they can be successfully formalised within first-order logic expressions over semantically rich building information models.

### 2.2. The dynamic nature of shadows and sunlight

There is an inherent dynamism in the nature of sunlight that operates at numerous levels. The location of the sun in relation to the environment clearly has an immediate impact on the forms of shadows. Although this relationship is constantly changing throughout the day, the induced patterns of sunlight and shadow can *appear* fixed at a given moment,

especially at an urban scale. Thus, an environment can exhibit distinct visual states from sunlight at various times of the day.

The seasonal variations of the sun's trajectory across the sky also has a clear impact in many countries: the low arc of the sun's path during winter can cast long shadows throughout a day and may leave some parts of the environment without any direct sunlight for months.<sup>2</sup>

This dynamism gives rise to two basic questions: what are the shapes of shadows and sunlight *at a given moment*, and what are those shapes *across periods of time*? For example, the designer may need to know whether a given room *ever* has direct sunlight during the day, or whether a certain cafe zone has direct sunlight at each midday period throughout the year.

Computationally answering each one of these questions requires the capability to handle conceptual, spatial, and temporal concepts in an expressive knowledge processing and inferencing environment. From an application viewpoint, this is the objective of the ongoing work reported in this paper.

### 3. A Qualitative Model of Sunlight and Shadows

In this section we describe the abstractions of our sun model. Our model provides the necessary information for determining the regions of space covered in shadow, it is flexible enough to work with both highly under-specified early designs and detailed numerical designs, and generates shadows in real-time. Thus, our model allows the designer to quickly experiment with a large number of designs to determine how shadows broadly behave.

In this model we focus explicitly on direct sunlight (and ignore ambient illumination [Schultz et al., 2009]), i.e. regions of space in which there is an uninterrupted straight line from the sun to every point in the region; shadows are the absence of direct sunlight.

#### 3.1. Modelling Outdoor Sunlight

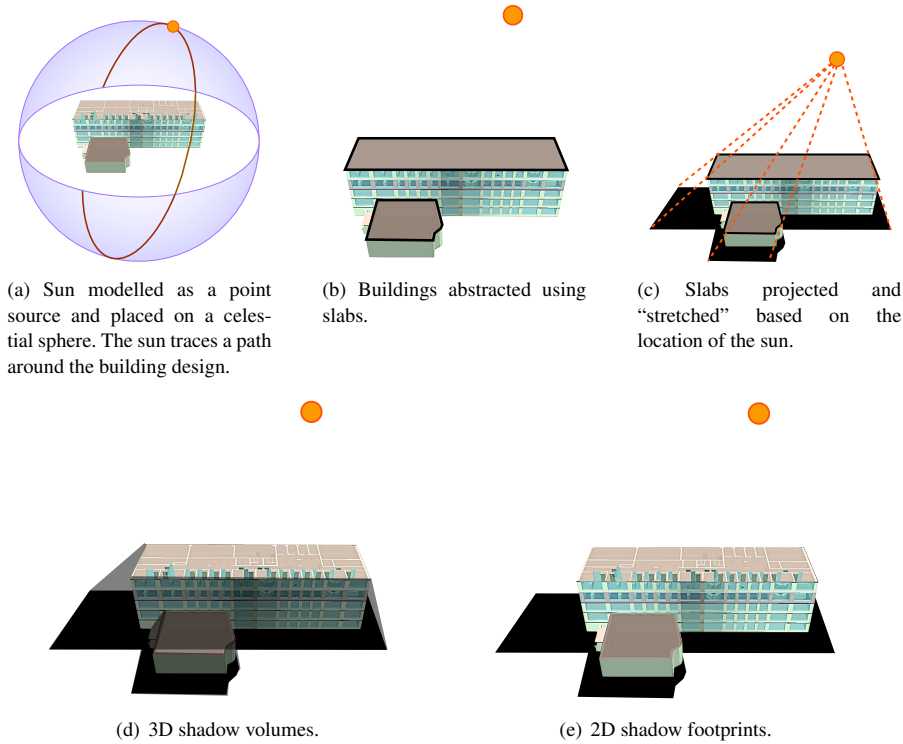
When modelling sunlight on an urban scale we only take large objects into account such as buildings, large trees, and billboards.<sup>3</sup> We use a highly abstracted model of buildings to determine how they cast outdoor shadows: while buildings can consist of thousands of objects, the single most informative features are the geometries, elevations, and heights of *slabs* (i.e. rooves and floors).

Shadows are generated from slabs by making the qualitative generalisation that walls and other objects holding up slabs are *effectively opaque*. This provides a rapid approximation of the *building envelope*. Other large outdoor objects are modelled using very abstract geometries such as bounding boxes. Thus, the designer only needs to provide a very rough outline of buildings and amenities to begin experimenting with shadows.

---

<sup>2</sup>The interference of other objects such as clouds can also influence the form of shadows, but these external conditions are not relevant for our qualitative model.

<sup>3</sup>Modelling the shadows of an outdoor park bench is too fine-grained and will not contribute to an understanding of how shadows influence and shape the environment on a large scale.



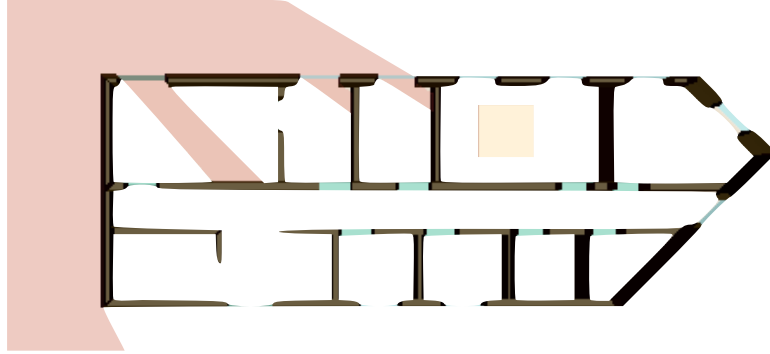
**Figure 2. Generating shadows with a qualitative sun model.**

As is typical for planetary models [Roderick, 1992], we model the sun as a point source placed on a celestial sphere centred on the design. The sun is positioned using altitude and azimuth angles.<sup>4</sup> We generate a model of shadows by “stretching” each vertex of each slab polygon according to the elevation and height of the slab, and location of the sun. This results in 2D shadow *footprints* (useful for top-down plan analysis) and 3D shadow *volumes*. The process is illustrated in Figure 2.

If the sun’s altitude is within a certain range then indoor direct sunlight can be approximated using the isovist of a point derived from the sun’s location,<sup>5</sup> as illustrate in Figure 3. This abstracts from the height information of the windows, and other precise design parameters, to provide the designer with information on the rooms that are exposed to direct sunlight.

<sup>4</sup>Standard formulae for converting latitude, longitude, calendar date, and time of day into approximate sun coordinates are usable; for details, refer to [Ast, 1996] page C24.

<sup>5</sup>More specifically, the sun point is projected onto a plane that is parallel to the ground and set to the elevation of the relevant building floor. The isovist is then taken from this projected point in that plane.



**Figure 3.** Direct sunlight in the building interior (the sun is located in the upper left area of the plan).

#### 4. Declarative High-level Design Specifications for Shadows

In this section we present an assortment of declarative programming logic specifications in the domain of urban and indoor design for art galleries, academic sites, and other public spaces. These example rules are processed using our general purpose spatial reasoning framework CLP(QS) [Bhatt et al., 2011]; CLP(QS) provides domain-independent geometric and qualitative spatial and temporal reasoning functionality within a Constraint Logic Programming (CLP) setting .

**Sunlight and Shadow predicates.** The location of the sun is defined by altitude and azimuth, which in turn can be computed from a calendar date and time. We can determine the sun location based on a given time, or the time that corresponds to a given location, by using the predicate:

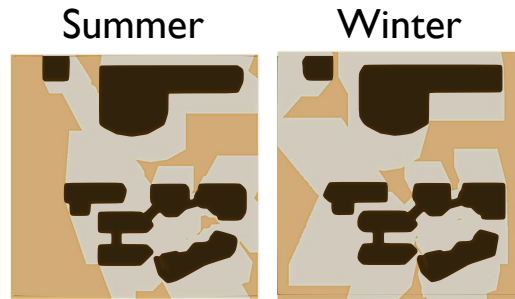
```
sun(DateTime, Azi, Alt).
```

**Painting damage.** There may be a risk that paintings are damaged if they are exposed to direct sunlight. This occurs if there is *some* period in which the sunlight overlaps the painting.

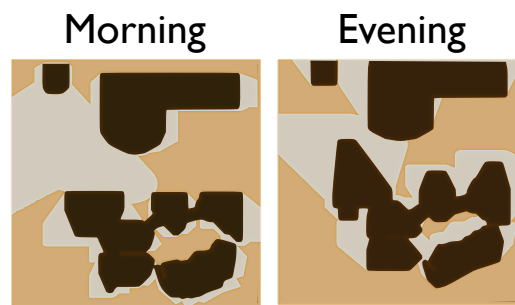
```
riskOfSunDamage(Painting) :-  
    physical_space(Painting, _, PPolygon),  
    sunlight(sun(DateTime, _, _), _, SPolygon),  
    topology(SPolygon, PPolygon, overlaps).
```

**Rooms with sunlight.** We can determine whether particular rooms in our design get direct sunlight during winter.

```
sunRoom(Room) :-  
    movement_space(_, Room, MPolygon),  
    winter(DateTime),  
    sunlight(sun(DateTime, _, _), _, SPolygon),  
    topology(SPolygon, MPolygon, overlaps).
```



**Figure 4.** Regions of space that are always in shadow (brown), and always in sunlight (orange), throughout the entire day, during summer and winter.



**Figure 5.** Regions of space that are always in shadow (brown) and always in sunlight (orange) for every morning (left) and evening (right) across the entire year.

**Uncomfortable cafe garden.** At around midday throughout the year, the area where people relax and enjoy lunch should not be completely covered in shadow.

```
uncomfortableCafeArea(CafeOutdoorArea) :-
    movement_space(CafeOutdoorArea,_,MPolygon),
    midday(DateTime),
    shadow(sun(DateTime,_,_),_,SPolygon),
    topology(SPolygon,MPolygon,contains).
```

**Shadow forms across periods of time.** The dynamism of shadows as a product of the changing relationship between environment and sun results in patterns and forms that can be captured and analysed. For example, as in the *uncomfortable cafe* example, the architect may need to study the topology of regions that are always in shadow over a period of time.

Figure 4 shows the areas throughout summer (left) and winter (right) that are always in sunlight (orange), and always in shadow (brown). We can observe that, during winter, certain areas of sunlight in the central portion of the design grow and become disconnected from the larger sunlit region to the right. This “island” of sunlight is highlighted

by the surrounding colder shadowed regions, and thus could be utilised as an area for outdoor winter activities such as winter markets, events, or a shared cafe area.

The periods of time do not need to be temporally contiguous. Figure 5 shows the areas of sunlight and shadow from every morning and evening throughout the entire year.

## 5. Discussion and Future work

In our current prototypical implementation, we compute dynamic forms of sunlight and shadow regions by sampling sun locations within the specified period, generating the sunlight or shadow regions for each sample, and combining the results. One future research aim is to fully encode the *relationship* between calendar date, sun location, and sunlight and shadow regions within the framework of Constraint Logic Programming. Furthermore, based on our preliminary foundations in [Schultz et al., 2009], we are also investigating the incorporation of non-uniformly lit ambient lighting conditions, which are more suited for interior design scenarios, and in cases where the internal composition of a building is controlled via artificial light sources.

This leads into the next core future research aim of generating *metric instantiations* from qualitative (spatial) relational specifications – that is, automatically adjusting designs in order to satisfy certain high-level qualitative properties. Given certain shadow properties that must be satisfied in the design such as “the painting must not be exposed to direct sunlight”, our reasoning system (i.e., underlying constraint solver) will find solutions by rotating and translating the objects on the ground-plane in a manner that the derived transformation satisfies the desired topological relationships between the physical entities and their cast shadows.

## References

- The Astronomical Almanac For the Year 1996*. Nautical Almanac Office, United States Naval Observatory and Her Majesty’s Nautical Almanac Office, 1996.
- M. Bhatt, J. H. Lee, and C. Schultz. CLP(QS): A Declarative Spatial Reasoning Framework. In *COSIT*, pages 210–230, 2011.
- M. Bhatt, C. Schultz, and M. Huang. The Shape of Empty Space: Human-centred Cognitive Foundations in Computing for Spatial Design. In *IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 33–40, 2012.
- M. L. Roderick. *Methods for calculating solar position and day length including computer programs and subroutines*. Land Management, Western Australian Department of Agriculture, 1992.
- C. P. L. Schultz, R. Amor, B. Lobb, and H. W. Guesgen. Qualitative design support for engineering and architecture. *Advanced Engineering Informatics*, 23:68–80, 2009.
- H. Seidler. Sunlight and architecture. *Architectural Science Review*, 2, 1959.

# Mental “structures” in the Berlin school of Gestalt Psychology: can sensation be described as “structural”?

Eric TREMAULT

*Phico/Execo, Université Paris 1 Panthéon-Sorbonne*

**Abstract.** It is not exaggerated to affirm that the modern notion of structure arises in Koffka's *Growth of the Mind* and in his following article, “Perception : An introduction to the Gestalt-theorie” (1922). The importance of the notion of structure as Koffka uses it lies in the fact that it is designed to replace the old empiricist notion of “sensation” as a real and separable element of the phenomenal field, corresponding to a definite stimulus. But, yielding to many suggestions by Köhler, Koffka does not only understand the interdependency of sensations in a structure as a causal one: in fact, he decidedly understands it as a logical one. Thus he defines structures as “very elementary reactions, which phenomenally are not composed of constituent elements, their members being what they are by virtue of their ‘member-character,’ their place in the whole; their essential nature being derived from the whole whose members they are” (“Perception”, p.543). I mean to show that the parts in such structures can only be what it is classical to name “relational attributes” or “relational predicates”. In other words, structures are now internal relations between their terms, and more precisely still “directly constitutive internal relations”, not internal relations reducing to the existence of their terms as were the internal relations against which Russell struggled, but relations to which their terms reduce. But the real importance of this notion of structure is that it rests and is built upon a truly impressive amount of empirical data. Nevertheless, I want to show that Koffka's conception of sensation is fundamentally impossible to conceive, and that the belief that it is empirically grounded rests mainly on a confusion between abstraction of a sense-datum and real separation of the stimuli underlying such a datum. As a consequence, phenomenal structures, if they exist, can only be external to their terms, as they are in Köhler's view, in spite of many ambiguities in his formulations. However, I will end by showing that, correctly understood, the notion of structure can still be of great help in phenomenology and psychology since it provides a naturalistic means to understand how a non-intentional “meaning” can be passively present at a sensory level.

**Keywords:** Structure ; Internal relations ; Sensation ; Gestalt psychology ; Phenomenology

## Introduction

My main aim in this paper is to object to any kind of “structural” theory concerning sensation, meaning by this any theory which claims that one cannot attribute a relation to a sensorial content without intrinsically altering it. In philosophical terms, such a theory would refuse to acknowledge a distinction between “knowledge by acquaintance” and “knowledge about”: it would argue that all so-called “knowledge by



acquaintance” is only knowledge so far as it is knowledge of the *relations* of a sensorial content, and that this content itself is ultimately reducible to those relations. Thus, a “structural theory of sensation” is essentially emphasizing the primacy of relations over sensations. Such was the neo-Hegelian approach to sensation at the end of the Nineteenth Century, whose most famous representatives were T.H. Green and F.H. Bradley. As William James put it in his *Principles of Psychology* ([1],[2]), quoting T.H. Green:

“The only reals for the neo-Hegelian writers appear to be *relations*, relations without terms, or whose terms are speciously such and really consist in knots, or gnarls relations finer still *in infinitum*. ‘Exclude from what we have considered real all qualities constituted by relation, we find that none are left’ ‘Abstract the many relations from the one thing and there is nothing ... Without relations it would not exist at all.’ [T. H. Green, *Prolegomena to Ethics*, §§ 20, 28.] ‘The single feeling is nothing real.’ ‘On the recognition of relations as constituting the *nature of ideas*, rests the possibility of any tenable theory of their reality.’ [Introduction to *Hume*, §§ 146, 188.]” ([2], p. 10).

Thus, it is T.H. Green who first developed most clearly a “structural theory” concerning sensation. Bradley’s point of view is more complex (as was particularly well shown by Peter Hylton [3]), since he also tried to reduce all external relations to internal relations, but then wanted to prove the unreality of internal relations themselves: for what he really tried to demonstrate was the unreality of *all* kinds of relations. Following the path of Russell and William James most notoriously, I shall then reject Bradley’s first move, and defend external relations against their reduction to internal ones; but then I shall nonetheless take over the arguments Bradley uses in his second move, against certain kinds of internal relations.

However, it is less known, and more directly interesting to us here, that the Berlin school of Gestalt psychology, when it first introduced the notion of “structure” in its contemporary meaning in psychology and philosophy, was itself mainly concerned with developing such a structural theory of sensation, in the neo-Hegelian sense just used. But it seems nevertheless that Gestalt psychologists never had any kind of *a priori* bias or leaning towards neo-Hegelian thinking when developing this notion of “structure” as it is still used today, and on the contrary always showed public disdain for what they called “romantic” theories of nature<sup>1</sup>. Even this disdain itself was only formulated as an answer to psychologists who precisely read Gestalt psychology as a new kind of neo-Hegelianism. Hence, there is no reason to think that this disdain was simulated, and more reason to think that Gestalt psychologists only discovered the possibility of tracing their theories back to neo-Hegelianism when reading the commentaries of others. Indeed, it is mainly Kurt Koffka who, among Berlin Gestalt psychologists, clearly developed a structural theory of sensations in *Die Grundlagen der psychischen Entwicklung* (1921, translated as [7] in 1924, first edition), and in his following article in English, “Perception : An introduction to the Gestalt-theorie” (1922, [8]), but this structural theory almost entirely disappeared in his late master work, *Principles of Gestalt Psychology* (1935, [9]), which largely explains why this aspect of Gestalt psychology seems to be widely ignored today, even though it remains as one of the central reasons for the influence of this school, especially among philosophers such

<sup>1</sup> See Köhler [4], pp. 153 f. ([5], p. 30), and especially his detailed answer to G.E. Müller on this question in [6] ([5], pp. 379 f.).

as Merleau-Ponty, Cassirer and Scheler. Now, as far as I know, Koffka never clearly explained why he had to abandon his initial theory, but the fact that it finally appeared to him linked with neo-Hegelianism may very well be one of the reasons for this disappearance. Nevertheless, I want to show that there are other reasons which go much deeper than this one. But what I wanted first to emphasize is that, not being interested at all in neo-Hegelian thought, the only reason why Gestalttheory ended up formulating a new kind of structural theory of sensation is that such a theory seems at first glance to rest on *facts*. This is what makes it so difficult to untangle the intellectual situation surrounding such structural theories even today: for a great number of psychological facts seem to support them, even though, as I intend to show, those theories will prove to be logically impossible to conceive. In this paper, I want to focus on this theoretical and logical impossibility of the notion of “structure” as Koffka first introduced it, and as it is still used today when precisely employed. But, as it seems to me that the real importance of this notion is that it rests and is built upon a truly impressive amount of empirical data, I need to begin with a paradigmatic example that will reveal its *prima facie* legitimacy.

### 1. Empirical data in favor of a “structural theory of sensation”

For this purpose, I will quickly present the phenomenon known since Jaensch as the phenomenon of “colour-transformation”, a phenomenon that Koffka lays great stress on since he uses it to interpret the chromatic constancy phenomena as well<sup>2</sup>. Koffka thus shows that all colours appear and are qualitatively determined upon a general “chromatic level” which may correspond to any objective colour stimulus but always appears as a neutral white, while the phenomenal colour of the other stimuli (which generally appear as “figures” upon this “ground”) depends upon their difference or “gradient” from the “level” stimuli. As a matter of fact, these level stimuli generally correspond approximately to the center of the chromatic scale of the present stimuli, and it is the *relation* of the surrounding stimuli to this chromatic center that determines the phenomenal colour those stimuli will appear with. Most of the time, the chromatic center is the chromatic value of the general *lighting*, so that the colour of this lighting will tend to phenomenally disappear, while the phenomenal colour of all stimuli will depend on their objective difference from the lighting stimuli. This explains why the figures we actually see always tend to appear with the same colour they “truly” have, even when seen under coloured lightings: inasmuch as the lighting covers both figure and ground, the “gradient” between them remains constant, whatever the objective colour of the lighting may be. But the “colour-transformation” phenomenon is most striking when the light on the figure is isolated, while the lighting on the ground is slowly changed. For, this time, the objective difference between the stimuli corresponding to figure and ground is accordingly modified, and the result is that the ground remains phenomenally constant while the colour of the figure changes under our eyes. For instance, if a “white” figure seen under neutral light is isolated from its “white” ground seen under a yellow light, the eye gets a neutral stimulus upon a “yellow” level, but the objective difference between figure and level will then seem to be “translated” so that what will actually appear is a “blue” figure upon a white “normal” ground. Thus, this phenomenon is extremely paradoxical if one still believes

---

<sup>2</sup> See Koffka, [8], pp. 567-570; [10], pp. 334 f.; and [9], pp. 254 f.

in what Köhler called “the constancy hypothesis” [10], i.e. the hypothesis that what really appears to us is in a continuous relation with the outside stimuli, so that it should be isomorphic to them. For here, the phenomenal figure changes from white to blue while its underlying stimuli remain constant, whereas the phenomenal ground remains white while its underlying stimuli change from white to yellow. The facts upon which Koffka built his first structural theory are thus generally facts that contradict the constancy hypothesis, and in which the phenomenal changes seem grounded on a change in the relations between the stimuli. Some of those facts, such as “contrast” phenomena, in which two adjacent colours in space or time tend to tinge with the complementary colour of each other, were known long before Gestalt psychology. But most of the empirical data used by Koffka to legitimate his structural conception of sensation is gathered from child and animal psychology. Koffka thus tries to show at great length in *The Growth of the Mind* [7] that children have to *learn to differentiate* colours and that only when they have begun to do so, do colours indeed appear to them. According to the detailed interpretation Koffka makes of the facts available at the time, even “things” or “figures” do not at first appear as such to children or animals, but only as members in more comprehensive “structures” which are more immediate or instinctive to children and animals than the “thing-structure” is. But let us now turn to those structures from a theoretical point of view, and begin our examination by clearly stating the way Koffka defines them.

## 2. Koffka’s definition of “structures”

The notion of mental structure first appears in Gestalt psychology in Köhler’s book on *Physical Gestalten* [4] but it is mainly developed by Koffka in *Die Grundlagen der psychischen Entwicklung* (1921), though in the English version of this book (*The Growth of the Mind* [7]), “*Struktur*” is translated as “configuration”, so that it won’t be confused with the notion of “structure” as used by Titchener at the same time<sup>3</sup>. But Titchener’s use of the word is no longer predominant, and as a matter of fact “structuralism” as we mean it today is largely based on the notion of structure as Koffka developed it<sup>4</sup>. So that it is not exaggerated to affirm that the modern notion of structure arises in Koffka’s *Growth of the Mind* and in his following article, “Perception : An introduction to the Gestalt-theorie” (1922, [8]), by which Koffka first

---

<sup>3</sup> “The translation of the book ... was a difficult task because of the new terminology employed, for which English equivalents had to be coined. The difficulty was increased by the fact that one of the chief terms employed, namely, *Struktur*, could not be retained as ‘structure’, since, as a result of the controversy between *structuralism* and *functionalism*, this term has a very definite and quite different meaning in English and American psychology. For want of a better term, we have chosen to follow a suggestion originally made by Professor E. B. Titchener, and have translated *Struktur* as ‘configuration,’ although I can not say that it has completely satisfied me” ([7], pp. xv-xvi).

<sup>4</sup> Of course this is not the place to trace the history of “structuralism”. For converging views, see for instance Merleau-Ponty [12], pp. 102 f.; and [13], pp. 142 f. Jakobson himself ([14], p. 715) reminded “the assiduous attention which linguists of the two hemispheres paid to the progress of Gestalt psychology” during the development of structural linguistics.

introduced Gestalt-theorie in English, and where *Struktur* was still translated as “structure”.

The importance of the notion of structure as Koffka uses it lies in the fact that it is clearly designed to replace the old empiricist notion of “sensation” as a real and separable element of the phenomenal field, corresponding to a definite stimulus. In that sense, as Merleau-Ponty put it in *La structure du comportement*, “the theory of form ... tends to develop in a philosophy of form which would substitute itself to the philosophy of substances” ([15], pp. 142-143). In *Physical Gestalten* ([4], p. 55; [5], p. 27), Köhler distinguished “structure” and “Gestalt” by admitting that “structures” are interdependent elements, thus revealing properties which they would not have, were they simply added one to another as in a pure “distribution”. But those properties are not only whole properties, and “structures” are not only wholes distinct from the sum of their parts, as Ehrenfels’ *Gestaltqualitäten* were ([4], pp. 35-27; [5], pp. 24-25): their very parts themselves are transformed by the structural nature of the whole. This amounts to saying that the whole which unites the parts is not only “formal”: it is a causal, dynamic whole, which is what a “Gestalt” means for the Berlin school of Gestalt Psychology. In other words, a mental “structure” between sensations reveals that there is a physiological “Gestalt” between their physiological correlates in the cortex. This is what the “isomorphism” hypothesis introduced by Gestalt psychologists means: each mental “form” (or “Gestalt quality”) is a “structure”, thus corresponding to a dynamic system in the brain. To talk about “structure” is to talk about the interdependent parts of this system, whether psychological or physiological; to talk about “Gestalt” for the Berlin school is to talk about the causal and dynamic whole which makes them qualitatively interdependent. But, yielding to many suggestions by Köhler, Koffka does not only understand the interdependency of sensations in a structure as a causal one: in fact, he decidedly understands it as a logical one. Here is how he defines “structures” in “Perception”:

“Structures, then, are very elementary reactions, which phenomenally are not composed of constituent elements, their members being what they are by virtue of their ‘member-character,’ their place in the whole; their essential nature being derived from the whole whose members they are” ([8], p.543).

Thus, Koffka wants to show that there are no absolute sensorial contents in our perception, but only structures. He takes the example of two squares of gray cardboard lying side by side, which we perceive to be of different grayness: can we describe this experience, as Ehrenfels and the Graz school would have done ([8], p. 536), as grounded on a comparison between two sets of otherwise atomic and independent sensations? In reality, Koffka says, what appears in this case is at once a differential structure, with a “steep or moderate ascent” ([8], p. 540), in one way or another, between the two squares. Hence, those do not appear for themselves, in isolation from each other, as two sets of sensations should, but they only appear as “steps” in a brightness scale:

“This must be rightly understood. If I say a real stair has two steps, I do not say there is one plank below and another plank above. I may find out later that the steps are planks, but originally I saw no planks, but only steps. Just so in my brightness steps: I see the darker left and the brighter right not as separate and independent pieces of color, but as steps, and as steps ascending from left to right. What does this mean? A plank is a plank anywhere and in any position; a step is a step only in its proper

position in a scale. Again, a sensation of gray, for traditional psychology, may be a sensation of gray anywhere, but a gray step is a gray step only in a series of brightnesses.” ([8], p. 540)

What we see, according to Koffka, is again a “‘crescendo’ or ‘diminuendo’”, which is “an undivided whole” ([8], p. 546), though it may be articulated into two different moments or “steps”. The main point is thus that those *steps* are really inseparable from the crescendo itself, which does not hold *between* them. On the contrary, the steps only hold *within* the crescendo:

“For, speaking of ‘steps’ I mean not only two different levels, but the rise itself, the upward trend and direction, which is not a separate, flighty, transitional sensation, but a central property of this whole undivided experience. Undivided does not mean uniform, for an undivided experience may be articulated and it may involve an immense richness of detail, yet this detail does not make of it a sum of many experiences. The direction upward or downward under certain conditions, e. g., under brief exposure, may be the chief moment of the total experience; in extreme cases, this direction may be present and nothing else, the plank-character of the steps having entirely vanished.” ([8], p. 541)

Defined as they are by Koffka, it seems to me that “structures” must then be understood as networks of internal relations, and that is what Koffka himself sometimes writes:

“Two colours adjacent to each other are not perceived as two independent things, but as having an inner connection which is at the same time a factor determining the special qualities A and B themselves.” ([7], p. 221)

But one must be very careful here to understand those *structural* internal relations in their very precise meaning. To talk about internal relations *logically* means that the terms of this relation would be different, were they not in this relation. Thus, a change in internal relation logically “implies” a change in the terms. But, as François Clementz most notably has clearly shown ([16], [17]), this can have two different *ontological* meanings: either the relation is grounded on its terms; or the terms are grounded on their relation. In the first case, the change in the relation “supposes” a change in the terms; in the second case, the change in the relation “determines” a change in the terms. The first type of internal relation is the most commonly discussed. “Similarity” for example, is generally admitted as an internal relation inasmuch as two white things cannot cease to be “similar” unless at least one of the things ceases to be white. But such is precisely not the way Koffka understands similarity in the case of two similar *sensations*: it is then the “inner connection” between the two colours that *determines* “the special qualities A and B themselves”. In this case, we are then dealing, not with a “grounded” internal relation (which are the only kind of internal relations discussed by Russell in his debate with Bradley), but with a holistic type of relation, that François Clementz (and also John Bacon [18]) calls “directly constitutive” internal relations. Clementz interestingly points out that the question whether “there *really* are internal relations in this sense – which seems to be what the British Idealists had in mind when they claimed that all relations are internal – is open to dispute” ([17], p. 172), but he adds in a note that, even today:

“Many philosophers would probably accept that there might well be relations of that kind – notably ‘structural’ relations – holding between such varieties of abstract, formal or intensional entities as space points,

numbers, concepts or meanings, phenomenal colours, social institutions, artworks and so on. Whether there are constitutive relations beyond this abstract domain is much more controversial. A widespread argument to the effect that they are no such relations obtaining between concrete particulars is that this would violate Hume's principle that there cannot be any kind of logical link between 'distinct existences'." ([17], p. 172 note 7)

It is worth noting here that, according to Clementz, phenomenal colours today only appear reducible to their relations on the condition that they are understood as "abstract, formal or intensional entities": but the question is precisely whether they are such, and I now want to show that they are not, beginning with "the widespread argument" that Clementz talks about.

### 3. Objections against Koffka's "structural programme"

Once Koffka's "structures" are understood as networks of "directly constitutive internal relations", it appears that some classical objections have been formulated against them, of which Koffka takes no account.

#### 3.1. "Structural relations" as inauthentic relations

The "widespread argument" Clementz refers to is indeed Bradley's classical objection against internal relations in *Appearance and Reality*, an argument very well summed up by Hylton:

"If *a* is internally related to *b*, then the relation to *b* is part of *a*'s internal nature. Since '*a*'s internal nature' is just what *a* essentially *is*, it follows that *a* is not independent, but is what it is only because of its relation to *b*. Internal relations are thus unstable: as relations they set up their objects as independent entities; as internal they make it clear that their objects are not independent, but can be considered only as part of a larger totality ... By their internality, internal relations make it manifest that they are destined to be transcended in a higher unity in which the separateness of the relata, and thus the relational nature of the whole, has disappeared." ([3], p. 55).

Interpreted most faithfully, this argument by Bradley seems to lead to the conclusion that internal relations cannot be *authentic* "relations", since they simply *cannot have any term*. Indeed, the ultimate goal of neo-Hegelian Idealism seems to be the reduction of all separate substances in traditional ontology to knots of "relational predicates", as a premise to demonstrate that only the "whole" uniting those "pseudo-substances" can be real. Thus, admitting that it is the *definition* of a relation to have terms, the pseudo-reality of internal relations is "destined to be transcended in a higher unity in which the separateness of the relata, and thus the relational nature of the whole, has disappeared". However, this argument thus formulated does not really bear against Koffka, since this "transcendence" of relations *with terms* actually seems to be what "structures" are destined to accomplish as well for Koffka, at least concerning sensations. The left square appears as the "less bright", the right one as "the brighter" of the two, and this *difference* in brightness is supposed to be *constitutive* of the brightness itself of both squares. But, if this is true, one important conclusion has to be drawn from the rejection

of Bradley's argument: the "steps" by which Koffka is trying to replace the old-fashioned separate sensations are nothing but what it is more traditional to call "relational predicates" – I shall rather say here "structural predicates". As such, those steps truly cannot be separated from the relations in which they are involved, since, as Russell said in his *Principles of Mathematics* ([19], § 214, p. 222), they are nothing but "cumbrous ways" of talking about relations (or structures) themselves.

### 3.2. *The need of an absolute ground for structures*

Nevertheless, it seems at first glance quite difficult to admit, and to conceive, that a brightness difference might be constitutive of two brightnesses. As a matter of fact, this is the central point around which this whole discussion revolves. A first obvious objection is anticipated by Koffka: isn't it obvious that a brightness difference has to be grounded on two different brightnesses? But the "isomorphism" hypothesis introduced by Gestalt psychology actually provides Koffka with a very easy and interesting answer to this objection: according to this hypothesis, phenomenal "structures" are supposed to be the immediate correlates of causal relations in the brain, those causal relations being supposed to hold between physiological processes linearly issued from stimuli. Thus, phenomenal "structures" are ultimately grounded on non phenomenal stimuli, and not on sensations:

"Here the argument may be anticipated that, in the analysis, parts must determine the whole; you lay the lighter gray at the left and you have a different brightness gradation than when you lay it at the right! But what does this argument really prove? Remember, you must not substitute your sensations for your stimuli. If you are careful not to do this, your argument must be that the arrangement of the single stimuli determines the whole structure. But you have not proved that the part phenomena have determined the whole phenomenon." ([8], pp. 543-544)

Therefore, there seems to be no contradiction in the psychological possibility that structures might appear without visible grounds. However, we still have to understand how plain "steps" inside those structures can finally appear, or seem to appear, as absolute qualities.

### 3.3. *Empirical refutation*

Now, the most radical idea in Koffka's "structural" programme (and probably the most radical idea in the Gestalt programme in general) is that to see a figure on a ground (and, hence, to see a "sensation" in the classical sense of the "mosaic" theory), is still to see a "structure" in Koffka's sense. This particular structure, of which figure and ground are thus only "steps", is called a "segregation" structure by the Gestalt psychologists. Thereby, Koffka writes in *The Growth of the Mind*, that "it is ... a part of the nature of a quality that it should lie upon a ground, or, as we may also say, that it should rise upon a level" ([7], p. 131). All "things" or "figures" we see are thus reduced to steps in segregation structures by Koffka. And he conversely maintains another very strong claim, according to which the ground itself phenomenologically depends upon such a "segregation" structure: therefore "mere ground would be equivalent to no consciousness at all" ([8], p. 566), so that "the most primitive phenomenon of consciousness is not the inarticulate ground-work, but the [structure (*configuration*)] or quality, which arises from this uniform background" ([7], p. 136). Merleau-Ponty in

particular has presented this claim as the center of the whole Gestalt theory in his *Introduction to Phénoménologie de la perception*, when he wrote that “the Gestalttheorie tells us that a figure on a ground is the most simple datum we can get”, so that “no point can be seen except as a figure on a ground”, and that “a truly homogeneous area, offering nothing to *perceive*, can be given to *no perception at all*” ([20], p. 26).

But it has to be stressed that meanwhile this last claim had been purely and simply refuted by Wolfgang Metzger, in a series of experiments published in 1930 [21], to which Koffka later devoted a central position in his *Principles of Gestalt Psychology* [9]. Indeed, Metzger managed to produce homogeneous stimulus conditions and observed that something *could* be perceived in those conditions: namely, “a mist of light which becomes more condensed at an indefinite distance” ([21], p. 13; quoted in [9], p. 111), and the whiteness of which is a function of the intensity of the light received. Although Koffka does not precisely state the problem, it is thus surely no coincidence that his initial claim that “the most primitive phenomenon of consciousness” is a segregation structure, is nowhere to be found in the *Principles*: it would clearly be in direct contradiction to Metzger’s results, since the correlate of this claim is that “mere ground would be equivalent to no consciousness at all”. On the contrary, it is now Metzger’s “mist of light” that Koffka establishes as “the simplest case” of perception (though this simplicity does not imply, it is true, any genetic primitivity, but only means a dynamic privilege, as the most “balanced” distribution):

“If perception is organization, i.e., a psychophysical process in extension depending upon the total stimulus distribution, then homogeneity of this distribution must be the simplest case and not the traditional one which contains a discontinuity.” ([9], p. 110)

By excluding here that any discontinuity in the stimulus distribution might produce a dynamically “simple” perception, it is not only the “traditional” case of a single sensation that Koffka henceforth considered as complex, but also any case of figure seen upon a ground. More importantly, by admitting, as he had to, that a pure ground can appear as a phenomenon, Koffka *ipso facto* ceased to consider this ground as a plain “step” in a segregation structure, which means he had to abandon his initial structural programme.

### 3.4. The “transposibility” of structures

Finally, at least two other *de jure* arguments can be objected to structural theories of sensation such as the one Koffka initially formalized. The first argument was precisely formulated by William James in his *Principles of Psychology* against the neo-Hegelian attempts to reduce sensation to relations. It was also directed against certain “sensationalist writers” such as Alexander Bain, who, on the basis of empirical data very similar to the ones later used by Koffka (e.g. the “contrast” phenomena mentioned above), “believe in a so-called ‘Relativity of Knowledge,’ which, if they only understood it, they would see to be identical with Professor Green’s doctrine. They tell us that the relation of sensations to each other is something belonging to their essence, and that no one of them has an absolute content” ([2], p. 11). James’ objection is particularly simple and effective: if all that was experienced, when listening to music, for instance, was the relations between the notes, one could not make any difference between two identical melodies played in different scales, since the relations between



the notes would be the same in both. Actually, one could not even tell the difference between any two pairs of notes:

“So far are we from not knowing (in the words of Professor Bain) ‘any one thing by itself, but only the difference between it and another thing,’ that if this were true the whole edifice of our knowledge would collapse. If all we felt were the *difference* between the C and D, or *c* and *d*, on the musical scale, that being the same in the pairs of notes, the pairs themselves would be the same, and language could get along without substantives.” ([2], p. 12)

I think it is fair to say that this argument anticipates the argument known as the “transposability” of Gestalt qualities, that Ehrenfels used the same year in his famous article [22]: since the relations between the sensorial contents can be transposed from one set of contents to another one, qualitatively different from the first, it is obvious that relations are something “more” than those contents. Ehrenfels used this argument against the reducibility of Gestalt qualities to their terms; James uses it against the reducibility of the terms to their relations. Of course, Koffka and Merleau-Ponty knew this argument by Ehrenfels: but they only referred to it through Köhler [4], who insisted on the necessity to admit that *physical* systems were Gestalt qualities too, since their whole properties could also be transposed from one set of physical substances to another. As a consequence, Koffka and Merleau-Ponty only spoke of the possibility to transpose structures from one set of *stimuli* or physiological processes to another, with the effect that the resulting phenomenal structure *and its phenomenal terms* remained the same despite the transposition. But if one insists on the fact that the transposition Ehrenfels himself talked about, as a criterion for Gestalt qualities, was a transposition from one set of *sensations* to another; then one immediately sees that this property of phenomenal Gestalt qualities as such is enough to refute Koffka’s initial structural programme.

### 3.5. Abstraction and real separation

I will only add one last argument against such a programme, which will help us understand why the facts so much seem to corroborate a structural theory of sensation. I will borrow this argument from Husserl’s mereology in his *Logical Investigations*, though one could also find the same general idea in James’ writings. This general idea is again quite simple: it very well might be that *in fact* no single stimulus can ever produce the same sensorial content in another context; it may even be that *in fact* no single sensorial content is ever the same for it constantly changes with the context in which it appears, which is itself in constant change. But the fact remains that *de jure*, it is always possible to consider such content for itself (a colour, for instance) and to *abstract* it from its context<sup>5</sup>:

---

<sup>5</sup> By this Husserl means that the content is “isolable in idea”, which precisely does not mean that “the actually *experienced* contents of the phenomenological sphere ... can be freed from all blending with coexistent contents”, but “means only that we can keep some content constant in idea despite boundless variation – variation that is free, though not excluded by a law rooted in the content’s essence – of the contents associated with it, and, in general, given with it. This means that it is unaffected by the elimination of any given arrangement of compresent contents whatsoever. This self-evidently entails: that the existence of this content, to the extent that this depends on

“In the ‘nature’ of the content itself, in its ideal essence, no dependence on other contents is rooted; the essence that makes it what it is, also leaves it unconcerned with all other contents. It may *as a matter of fact* be that, with the existence of this content, other contents are given, and in accordance with empirical rules. In its ideally graspable essence, however, the content is independent; this essence by itself, i.e. considered in *a priori* fashion, requires no other essence to be interwoven with it.” ([23] p. 9; see also pp. 6-7)

On the contrary, Husserl adds, it is *de jure* impossible to abstract a structural predicate, or what Husserl calls a “moment of unity” in an intuitive content, from the whole or form-quality of which it is a moment ([23], p. 8). Indeed, as we already said with Russell, such structural predicates are only “cumbrous” ways of talking about the structure itself. Thus, it seems to me that Husserl’s argument can be summed up this way: one must not confuse the *abstraction* that can always be made of a sensorial content whatsoever, with the *real separation* from its context of the stimulus beneath it. It is this confusion that made it seem possible to think that facts could support a structural theory of sensation. That the phenomenal *effects* of *stimuli* may depend on the context of their presentation does not imply that those effects are *in themselves* dependent on their phenomenal context. Even if those effects only existed for an instant, they would exist as absolute beings, whereas structural predicates can only be relative beings: therefore, the possibility to abstract colours can only be conceived if they are not relational predicates, and this proves that they have to be admitted as irreducible absolutes in our ontology. As a consequence, phenomenal structures or form-qualities must be conceived as external to their terms, even though they might still be accepted as immediate phenomena, according to the “isomorphism” hypothesis.

As a matter of fact, this last hypothesis makes it very easy to conceive why, *de facto*, almost no phenomenal change in the relations can ever occur without concomitant change in the quality of the phenomenal terms. If those *absolute* terms are themselves correlates of absolute physiological processes, the “isomorphism” hypothesis entails that those processes are causally interdependent, since they are phenomenally structured. Now, it is a truism to say that causal relations can modify their terms. It may very well be for instance that, in the “colour-transformation” phenomena, the effect of those causal relations on the “level” processes is always to transform them into objectively “white” processes, and that those same causal relations actually accordingly affect the surrounding colour processes, in such a way that the objective difference between the colour stimuli might be preserved and translated between the colour processes in the brain. Actually, such is the way Koffka finally understood those phenomena in his later works ([10], [9]; see also [24], pp. 232-234). The resultant phenomenal colours would immediately appear as “steps” in “difference” or “segregation” structures, so they would immediately be *felt as* having the meaning of being different from each other, but they would nevertheless be absolute in themselves, and those structures would remain external to them.

---

itself and its essence, is not at all conditioned by the existence of other contents, that it could exist as it is, through an *a priori* necessity of essence, even if nothing were there outside of it, even if all around were altered at will, i.e. without principle.” ([23], p. 9).

## References

- [1] W. James, *Principles of Psychology*, volume I, Henry Holt & Co, New York, 1890.
- [2] W. James, *Principles of Psychology*, volume II, Henry Holt & Co, New York, 1890.
- [3] P. Hylton, *Russell, Idealism and the Emergence of Analytic Philosophy*, Oxford University Press, USA, 1993.
- [4] W. Köhler, *Die physischen Gestalten in Ruhe und im stationären Zustand, Eine naturphilosophische Untersuchung*, Erlangen, 1920.
- [5] W. D. Ellis (ed.), *A Source Book of Gestalt Psychology*, Routledge & Kegan Paul, London, 1938 (The Gestalt Journal Press, Highland, New York, 1997).
- [6] W. Köhler, "Komplextheorie und Gestalttheorie, Antwort auf G. E. Müllers Schrift gleichen Namens", *Psychologische Forschung*, 6 (1925), 358-416.
- [7] K. Koffka, *The Growth of the mind, An introduction to Child-Psychology*, second edition, second impression, translated by M. R. Ogden, Harcourt, Brace & Co., New York, 1927 (Kessinger Publishing, 2007).
- [8] K. Koffka, "Perception: An introduction to the Gestalt-theorie", *Psychological Bulletin*, 19 (1922), 531-585.
- [9] K. Koffka, *Principles of Gestalt Psychology*, Harcourt, Brace & World, New York, 1935.
- [10] K. Koffka, "Some Remarks on the Theory of colour Constancy", *Psychologische Forschung*, 16 (1932), 329-354.
- [11] W. Köhler, "On unnoticed sensations and errors of judgment", *The selected papers of Wolfgang Köhler*, M. Henle (ed.), Liveright Publishing Corporation, New York, 1971, 13-39.
- [12] M. Merleau-Ponty, *Sens et non-sens*, Gallimard, NRF, Paris, 1996.
- [13] M. Merleau-Ponty, *Signes*, Gallimard, NRF, Paris, 1960.
- [14] R. Jakobson, "Retrospect", *Selected Writings*, 2. *Word and language*, Mouton, The Hague-Paris, 1971, 711-722.
- [15] M. Merleau-Ponty, *La structure du comportement*, PUF, Quadrige, Paris, 1990.
- [16] F. Clementz, "Réalité des relations et relations causales", *La structure du monde, objets, propriétés, états de chose, Renouveau de la métaphysique dans l'école australienne de philosophie*, Jean-Maurice Monnoyer (ed.), Vrin, Paris, 2004, 495-521.
- [17] F. Clementz, "Relational Truthmakers", *Metaphysics and Truthmakers*, J. M. Monnoyer (ed.), Ontos Verlag, Frankfurt, 2007, 163-198.
- [18] J. Bacon, *Universals and Property Instances: The Alphabet of Being*, Blackwell, Oxford, 1995.
- [19] B. Russell, *The Principles of Mathematics*, Cambridge University Press, Cambridge, 1903.
- [20] M. Merleau-Ponty, *Phénoménologie de la perception*, Gallimard, Tel, Paris, 2010 (new impression).
- [21] W. Metzger, "Optische Untersuchungen am Ganzfeld II. Zur Phänomenologie des homogenen Ganzfelds", *Psychologische Forschung*, 8 (1930), 6-29.
- [22] C. Ehrenfels, "Über 'Gestaltqualitäten'", *Vierteljahrsschrift für wissenschaftliche Philosophie*, XIV, 3 (1890), 249-292.
- [23] E. Husserl, *Logical investigations*, Vol. II, Part II, Routledge & Kegan Paul, International Library of Philosophy, 2001.
- [24] E. Trémault, *Structure et sensation dans la psychologie de la forme, chez Maurice Merleau-Ponty et William James*, Ph.D. dissertation under the supervision of Jocelyn Benoist, Université Paris 1 Panthéon Sorbonne, 2013; text available on line: [http://www.academia.edu/2449951/Structure\\_et\\_sensation\\_dans\\_la\\_psychologie\\_de\\_la\\_forme\\_chez\\_Maurice\\_Merleau-Ponty\\_et\\_William\\_James](http://www.academia.edu/2449951/Structure_et_sensation_dans_la_psychologie_de_la_forme_chez_Maurice_Merleau-Ponty_et_William_James)

# Statistical Invariants of Spatial Form: From Local AND to Numerosity

Christoph ZETZSCHE<sup>1</sup>, Konrad GADZICKI and Tobias KLUTH  
*Cognitive Neuroinformatics, University of Bremen, Germany*

**Abstract** Theories of the processing and representation of spatial form have to take into account recent results on the importance of holistic properties. Numerous experiments showed the importance of “set properties”, “ensemble representations” and “summary statistics”, ranging from the “gist of a scene” to something like “numerosity”. These results are sometimes difficult to interpret, since we do not exactly know how and on which level they can be computed by the neural machinery of the cortex. According to the standard model of a local-to-global neural hierarchy with a gradual increase of scale and complexity, the ensemble properties have to be regarded as high-level features. But empirical results indicate that many of them are primary perceptual properties and may thus be attributed to earlier processing stages. Here we investigate the prerequisites and the neurobiological plausibility for the computation of ensemble properties. We show that the cortex can easily compute common statistical functions, like a probability distribution function or an autocorrelation function, and that it can also compute abstract invariants, like the number of items in a set. These computations can be performed on fairly early levels and require only two well-accepted properties of cortical neurons, linear summation of afferent inputs and variants of nonlinear cortical gain control.

**Keywords.** shape invariants, peripheral vision, ensemble statistics, numerosity

## Introduction

Recent evidence shows that our representation of the world is essentially determined by holistic properties [1,2,3,4,5,6]. These properties are described as “set properties”, “ensemble properties”, or they are characterized as “summary statistics”. They reach from the average orientation of elements in a display [1] over the “gist of a scene” [7,8], to the “numerosity” of objects in a scene [9]. For many of these properties we do not exactly know by which kind of neural mechanisms and on which level of the cortex they are computed. According to the standard view of the cortical representation of shape, these properties have to be considered as high-level features because the cortex is organized in form of a local-to-global processing hierarchy in which features with increasing order of abstraction are computed in a progression of levels [10]. At the bottom, simple and locally restricted geometrical features are computed, whereas global and complex properties are represented at the top levels of the hierarchy. Across levels, invariance is system-

---

<sup>1</sup>Corresponding Author: Christoph Zetzsche, Cognitive Neuroinformatics, FB3, University of Bremen, P.O. Box 330 440, 28334 Bremen, Germany; E-mail: zetzsche@informatik.uni-bremen.de  
Research supported by DFG (SFB/TR 8 Spatial Cognition, A5-[ActionSpace])

atically increased such that the final stages are independent of translations, rotations, size changes, and other transformations of the input. However convincing this view seems on first sight, it creates some conceptual difficulties.

The major difficulty concerns the question of what exactly is a low-level and a high-level property. Gestalt theorists already claimed that features considered high-level according to a structuralistic view are primary and basic in terms of perception. Further doubts have been raised by global precedence effects [11]. Similar problems arise with the recently discovered ensemble properties. The gist of a scene, a high-level feature according to the classical view, can be recognized in 150 msec [7,12,13,14] and can be modeled using low-level visual features [8]. In addition, categories can be shown to be faster processed than basic objects, contrary to the established view of the latter as entry-level representations [15]. A summary statistics approach, also based on low-level visual features, can explain the holistic processing properties in the periphery of the visual field [4,16,17]. What is additionally required in these models are statistical measures, like probability distributions and autocorrelation functions, from which it is not known how and on which level of the cortical hierarchy they can be realized.

One of the most abstract ensemble properties seems to be the number of elements in a spatial configuration. However, the ability to recognize this number is not restricted to humans with mature cognitive abilities but has also been found in infants and animals [9,18], recently even in invertebrates [19]. Neural reactions to numerosity are fast (100 msec in macaques [20]). And finally there is evidence for a “direct visual sense for number” since number seems to be a primary visual property like color, orientation or motion, to which the visual system can be adapted by prolonged viewing [21].

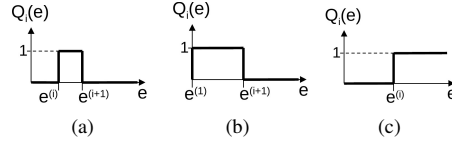
The above observations on ensemble properties raise a number of questions, from which the following are addressed in this paper: Sect. 1: Can the cortex compute a probability distribution? Sect. 2: And also an autocorrelation function? By which kind of neural hardware can this be achieved? Sect.3: Can the shape of individual objects also be characterized by such mechanisms? Sect. 4: What is necessary to compute such an abstract property like the number of elements in a spatial configuration? Can this be achieved in early sensory stages?

## 1. Neural Computation of a Probability Distribution

Formally, the probability density function  $p_e(e)$  of a random variable  $\mathbf{e}$  is defined via the cumulative distribution function:  $p_e(e) \triangleq \frac{dP_e(e)}{de}$  with  $P_e(e) = \Pr[\mathbf{e} \leq e]$ . Their empirical counterparts, the histogram and the cumulative histogram, are defined by use of *indicator functions*. For this we divide the real line into  $m$  bins  $(e^{(i)}, e^{(i+1)})$  with bin size  $\Delta e = e^{(i+1)} - e^{(i)}$ . For each bin  $i$ , an indicator function is defined as

$$Q_i(e) = 1_i(e) = \begin{cases} 1, & \text{if } e^{(i)} < e \leq e^{(i+1)} \\ 0, & \text{else} \end{cases} \quad (1)$$

An illustration of such a function is shown in Fig. 1a. From  $N$  samples  $e_k$  of the random variable  $\mathbf{e}$  we then obtain the histogram as  $h(i) = \frac{1}{N} \sum_{k=1}^N Q_i(e_k)$ . The cumulative histogram  $H_e(e)$  can be computed by changing the bins to  $(e^{(1)}, e^{(i+1)})$  (cf. Fig. 1b), and by performing the same summation as for the normal histogram. The reverse cumulative



**Figure 1.** Indicator functions. Basic types are: (a) indicator function for computation of a classical histogram. (b) indicator function for a cumulative histogram. (c) indicator function for a reverse cumulative histogram.

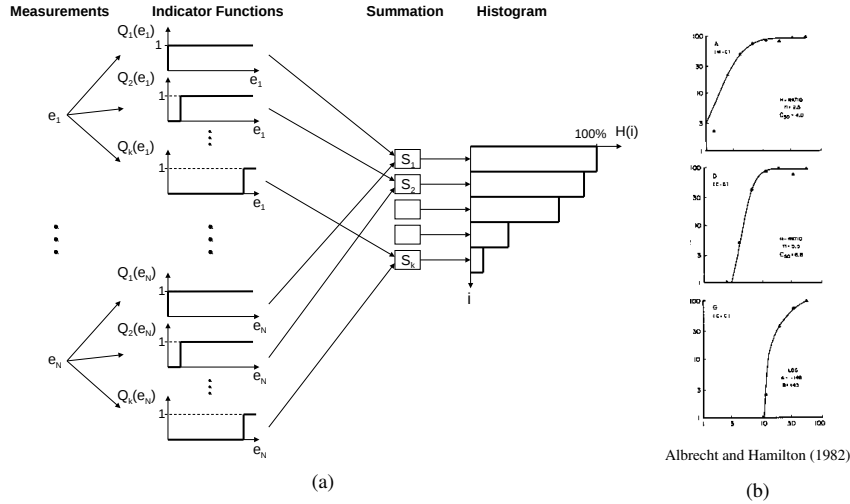
histogram  $\bar{H}(i)$  is simply the reversed version of the cumulative histogram. The corresponding bins are  $\Delta e_i = (e^{(i)}, e^{(m+1)}]$  and the indicator functions are defined as (Fig. 1c)

$$Q_i(e) = 1_i(e) = \begin{cases} 1, & \text{if } e \geq e^{(i)} \\ 0, & \text{else} \end{cases} \quad (2)$$

The corresponding system is shown in Fig. 2.

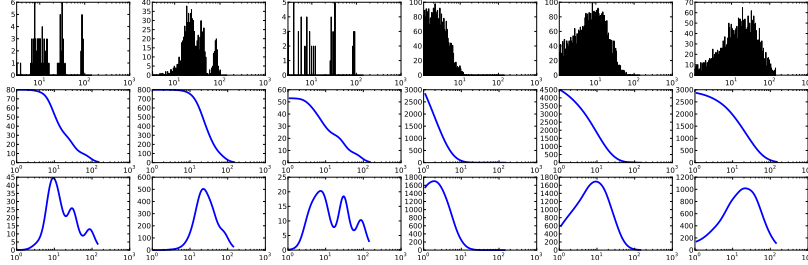
The three types of histograms have identical information content since they are related to each other as

$$h(i) = H((i+1)) - H(i) = \bar{H}(i) - \bar{H}(i+1) \quad \text{and} \quad H(i) = 1 - \bar{H}(i) = \sum_{j=1}^i h(j). \quad (3)$$



**Figure 2.** Computation of the reverse cumulative histogram. (a) shows the set of input variables  $e_1$  to  $e_n$  over which the histogram should be computed. Each of these variables is input to a set of *indicator functions*  $Q_i(e_k)$ . For each bin of the histogram there is a summation unit  $S_i$  which sums over all indicator function outputs with index  $i$ , i.e. over all  $Q_i(e_k)$ .

(b) The response functions of three neurons in the visual cortex [22]. They show a remarkable similarity to the indicator functions for the reverse cumulative histogram. First, they come with different sensitivities. Second, they exhibit an independence on the input strength: once the threshold and the following transition range is exceeded the output remains constant and does no longer increase when the input level is increased.



**Figure 3.** Neurobiological computation of a reverse cumulative histogram. The upper row shows several examples of input probability distributions. The second row shows the corresponding reverse cumulative histograms computed by a dense set of simulated neurons. The third row shows the estimated probability distributions as derived from the neural representation by use of Eq. (3).

How does all this relate to visual cortex? Has the architecture shown in Fig. 2a any neurobiological plausibility? The final summation stage is no problem since the most basic capability of neurons is computation of a linear sum of their inputs. But how about the indicator functions? They have two special properties: First, the indicator functions come with different *sensitivities*. An individual function does only generate a non-zero output if the input  $e$  exceeds a certain level, a kind of threshold, which determines the sensitivity of the element  $e^{(i)}$  in Eq. (2) and Fig. 1c. To cover the complete range of values, different functions with different sensitivities are needed (Fig. 2a). Second, the indicator functions exhibit a certain *independence of the input level*. Once the input is clearly larger than the threshold, the output remains constant (Fig. 1c).

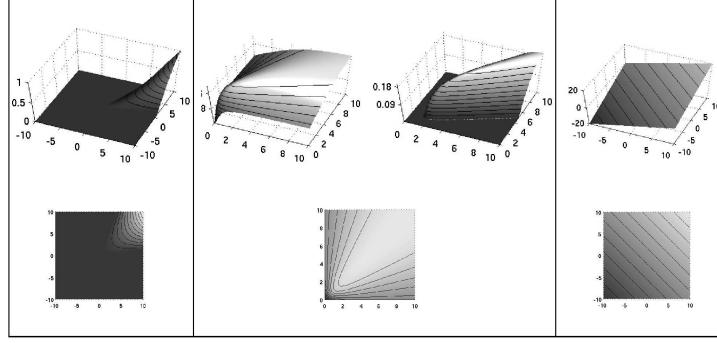
Do we know of neurons which have such properties, a range of different sensitivities, and a certain independence of the input strength? Indeed, cortical gain control (or normalization), as first described in early visual cortex (e.g. [22]) but now believed to exist throughout the brain [23], yields exactly these properties. Gain-controlled neurons (Fig. 2b) exhibit a remarkable similarity to the indicator functions used to compute the reverse cumulative histogram, since they (i) come with different sensitivities, and (ii) provide an independence of the input strength in certain response ranges.

The computation of a reverse cumulative histogram thus is well in reach of the cortex. We only have to modify the architecture of Fig. 2a by the smoother response functions of cortical neurons. The information about a probability distribution available to the visual cortex is illustrated in Fig. 3. The reconstructed distributions, as estimated from the neural reverse cumulative histograms, are a kind of Parzen-windowed (lowpass-filtered) versions of the original distributions.

## 2. Neural Implementation of Auto- and Cross-Correlation Functions

A key feature of the recent statistical summary approach to peripheral vision [4,6,24,16] is the usage of auto- and cross-correlation functions. These functions are defined as

$$h(i) = \frac{1}{N} \sum_{k=-N/2+1}^{N/2} e(k) \circ g(i+k), \quad (4)$$



**Figure 4.** Different types of AND-like functions. Each function is of the type  $g_k = g(s_i, s_j)$ , i.e. assigns an output value to each combination of the two input values. The upper row shows the functions as surface plots, the lower row as iso-response curves. Left: Mathematical multiplication of two inputs. Center: AND-like combinations that can be obtained by use of cortical gain control (normalization). The upper left figure shows the classical gain control without additional threshold. The upper right figure shows the same mechanism with an additional threshold. This results in a full-fledged AND with a definite zero response in case that only one of the two inputs is active. Right: The linear sum of the two input values for comparison purposes.

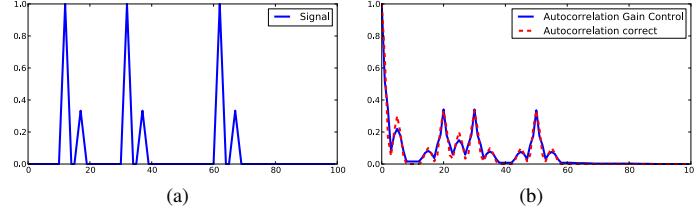
where autocorrelation results if  $e(k) = g(k)$  and where  $\circ$  indicates multiplication. With respect to their neural computation, the outer summation is no problem, but the crucial function is the *nonlinear multiplicative interaction between two variables*. A neural implementation could make use of the Babylonian trick  $ab = \frac{1}{4}[(a+b)^2 - (a-b)^2]$  [25,26,27], but this requires two or more neurons for the computation and thus far there is neither evidence for such a systematic pairing of neurons nor for actual multiplicative interactions in the visual cortex. However, exact multiplication is not the key factor: a reasonable statistical measure merely requires provision of a matching function such that  $e(k)$  and  $g(i+k)$  generate a large contribution to the autocorrelation function if they are similar, and a small contribution if they are dissimilar. For this, it is sufficient to provide a neural operation which is AND-like [27,28]. Surprisingly, such an AND-like operation can be achieved by the very same neural hardware as used before, the cortical gain control mechanism, as shown in [28]. Cortical gain control [22,29] applied to two different features  $s_i(x, y)$  and  $s_j(x, y)$  can be written as

$$g_k(x, y) = g(s_i(x, y), s_j(x, y)) := \max \left( 0, \frac{s_i + s_j}{(\sqrt{s_i^2 + s_j^2 + \epsilon})\sqrt{2}} - \Theta \right) \quad (5)$$

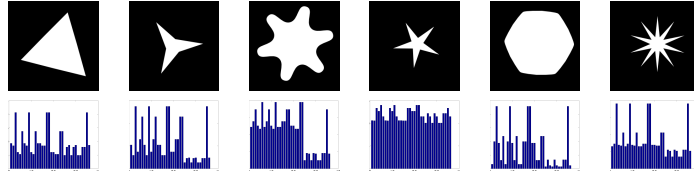
where  $k = k(i, j)$ ,  $\epsilon$  is a constant which controls the steepness of the response and  $\Theta$  is a threshold. The resulting nonlinear combination is comparable with an AND-like operation of two features and causes a substantial nonlinear increase of the neural selectivity, as illustrated in Fig. 4.

Of course there will be differences between a formal autocorrelation function and the neurobiological version, but the essential feature, the signaling of good matches in dependence of the relative shifts will be preserved (Fig. 5).





**Figure 5.** Mathematical and neurobiological autocorrelation functions. (a) shows a test input and (b) the corresponding mathematical (red dotted) and neurobiological (blue) autocorrelation function.



**Figure 6.** Different shapes and the corresponding integral features. We used parameter combinations of six different orientations  $\theta_i = (i-1)\pi/6$ ,  $i = 1, \dots, 6$ , and four different scales  $r_i = 2^{-i}$ ,  $i = 1, \dots, 4$ . The radial half-bandwidth was set to  $f_{r,h} = \frac{1}{3}r$  and the angular half-bandwidth was constant with  $f_{\theta,h} = \pi/12$ . Each parameter combination creates pairs of variables for each x,y-position which are AND-combined by the gain control mechanism described in Eq. (5) as  $g_k(x,y) = g(s_i(x,y), s_j(x,y))$ .

### 3. Figural Properties from Integrals

We extracted different features  $s_{r,\theta}$  from the image luminance function  $l = l(x,y)$  by applying a Gabor-like filter operation  $s_{r,\theta}(x,y) = (l * \mathcal{F}^{-1}(H_{r,\theta}))(x,y)$  where  $\mathcal{F}^{-1}$  denotes the inverse Fourier transformation and the filter kernel  $H_{r,\theta}$  is defined in the spectral space. We distinguish two cases (even and odd) which can be seen in the following definition in polar coordinates:

$$H_{r,\theta}^{even}(f_r, f_\theta) := \begin{cases} \cos^2\left(\frac{\pi}{2} \frac{f_r - r}{2f_{r,h}}\right) \cos^2\left(\frac{\pi}{2} \frac{f_\theta - \theta}{2f_{\theta,h}}\right) & , (f_r, f_\theta) \in \Omega_{r,\theta} \\ 0 & , \text{else,} \end{cases}$$

with  $\Omega_{r,\theta} := \{(f_r, f_\theta) | f_r \in [r - 2f_{r,h}, r + 2f_{r,h}] \wedge f_\theta \in [\theta - 2f_{\theta,h}, \theta + 2f_{\theta,h}] \cap [\theta + \pi - 2f_{\theta,h}, \theta + \pi + 2f_{\theta,h}]\}$ , where  $f_{r,h}$  denotes the half-bandwidth in radial direction and  $f_{\theta,h}$  denotes the half-bandwidth in angular direction.  $H_{r,\theta}^{odd}$  is defined as the Hilbert transformed even symmetric filter kernel.

Various AND combinations of these oriented features (see caption Fig. 6) are obtained by the gain-control mechanism described in Eq. (5). The integration over the whole domain results in *global* features  $F_k := \int_{\mathbb{R}^2} g_k(x,y) d(x,y)$  which capture basic shape properties (Fig. 6).

### 4. Numerosity and Topology

One of the most fundamental and abstract ensemble properties is the number of elements of a set. Recent evidence (see Introduction) raised the question at which cortical level

the underlying computations are performed. In this processing, a high degree of invariance has to be achieved, since numerosity can be recognized largely independent of other properties like size, shape and positioning of elements. Models which address this question in a neurobiologically plausible fashion, starting from individual pixels or neural receptors instead of an abstract type of input, are rare. To our knowledge, the first approach in this direction has been made in [30]. A widely known model [31] has a shape-invariant mapping to number which is based on linear DOG filters of different sizes, which substantially limits the invariance properties. A more recent model is based on unsupervised learning but has only employed moderate shape variations [32]. In [30] we suggested that the necessary invariance properties may be obtained by use of a theorem which connects local measurements of the differential geometry of the image surface with global topological properties [30,33]. In the following we will build upon this concept.

The key factor of our approach is a relation between surface properties and a topological invariant as described by the famous Gauss-Bonnet theorem. In order to apply this to the image luminance function  $l = l(x, y)$  we interpret this function as a surface  $S := \{(x, y, z) \in \mathbb{R}^3 | (x, y) \in \Omega, z = l(x, y)\}$  in three-dimensional real space. We then apply the formula for the Gaussian curvature

$$K(x, y) = \frac{l_{xx}(x, y)l_{yy}(x, y) - l_{xy}(x, y)^2}{(1 + l_x(x, y)^2 + l_y(x, y)^2)^2}, \quad (6)$$

where subscript denotes the differentiation in the respective direction (e.g.  $l_{xy} = \frac{\partial^2 l}{\partial x \partial y}$ ). The numerator of (6) can also be written as  $D = \lambda_1 \lambda_2$  where  $\lambda_{1,2}$  are the eigenvalues of the Hessian matrix of the luminance function  $l(x, y)$  which represent the partial second derivatives in the principal directions. The values and signs of the eigenvalues give us the information about the shape of the luminance surface  $S$  in each point, whether it is elliptic, hyperbolic, parabolic, or planar. Since Gaussian curvature results from the multiplication of the second derivatives  $\lambda_{1,2}$  it is zero for the latter two cases. It has been shown that this measure can be generalized in various ways, in particular towards the use of neurophysiologically realistic Gabor-like filters instead of the derivatives [27,30]. The crucial point, however, is the need for *AND combinations of oriented features* [27,30] which can be obtained as before by the neural mechanism of cortical gain control [28].

The following corollary from the Gauss-Bonnet theorem is the basis for the invariance properties in the context of numerosity.

**Corollary 4.1** *Let  $S \subset \mathbb{R}^3$  be a closed two-dimensional Riemannian manifold. Then*

$$\int_S K \, dA = 4\pi(1 - g) \quad (7)$$

*where  $K$  is the Gaussian curvature and  $g$  is the genus of the surface  $S$ .*

We consider the special case where the luminance function consists of multiple objects (polyhedra with orthogonal corners) with constant luminance level. We compare the surface of this luminance function to the surface of a cuboid with holes that are shaped like the polyhedra. The trick is that the latter surface has a genus which is determined by the number of holes in the cuboid and which can be determined by the integration of the local curvature according to Eq. (7). If we can find the corresponding contributions of

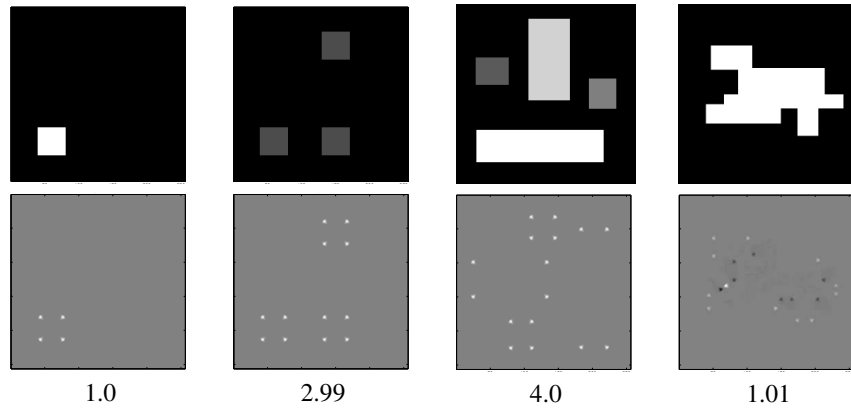
the integral on the image surface, we can use this integral to count the number of objects. We assume the corners to be locally sufficiently smooth such that the surfaces are Riemannian manifolds. The Gaussian curvature  $K$  then is zero almost everywhere except on the corners. We hence have to consider only the contributions of the corners. It turns out that these contributions can be computed from the elliptic regions only if we use different signs for upwards and downwards oriented elliptic regions. We thus introduce the following operator which distinguishes the different types of ellipticity in the luminance function. Let  $\lambda_1 \geq \lambda_2$ , then the operator  $N(x, y) := |\min(0, \lambda_1(x, y))| - |\max(0, \lambda_2(x, y))|$  is always zero if the surface is hyperbolic and has a positive sign for positive ellipticity and a negative one for negative ellipticity. We thus can calculate the numerosity feature which has the ability of counting objects in an image by counting the holes in an imaginary cuboid as follows:

$$F = \int_{\Omega} \frac{N(x, y)}{(1 + l_x(x, y)^2 + l_y(x, y)^2)^{\frac{3}{2}}} d(x, y). \quad (8)$$

The crucial feature of this measure are contributions of fixed size and with appropriate signs from the corners. The denominator can thus be replaced by a neural gain control mechanism and an appropriate renormalization. For the implementation here we use a shortcut which gives us straight access to the eigenvalues. The numerator  $D(x, y)$  of (6) can be rewritten as

$$D(x, y) = l_{xx}l_{yy} - \frac{1}{4}(l_{uu} - l_{vv})^2 = \frac{1}{4}[(l_{xx} + l_{yy})^2 - \underbrace{(l_{xx} - l_{yy})^2 + (l_{uu} - l_{vv})^2}_{=: \varepsilon^2}] = \frac{1}{4}(\Delta l^2 - \varepsilon^2) \quad (9)$$

with  $u := x \cos(\pi/4) + y \sin(\pi/4)$  and  $v := -x \sin(\pi/4) + y \cos(\pi/4)$ . The eigenvalues then are  $\lambda_{1,2} = \frac{1}{2}(\Delta l \pm |\varepsilon|)$  and we can directly use them to compute  $N(x, y)$ . Application of this computation to a number of test images is shown in Fig. 7.



**Figure 7.** Based on a close relation to topological invariants the spatial integration of local curvature features can yield highly invariant numerosity estimates. The numerical values in the last row are the normalized integrals of the filter outputs (middle row).

## 5. Conclusion

Recent evidence shows that ensemble properties play an important role in perception and cognition. In this paper, we have investigated by which neural operations and on which processing level statistical ensemble properties can be computed by the cortex. Computation of a probability distribution requires *indicator functions* with different sensitivities, and our reinterpretation of cortical gain control suggests that this could be a basic function of this neural mechanism. The second potential of cortical gain control is the computation of *AND-like feature combinations*. Together with the linear summation capabilities of neurons this enables the computation of powerful invariants and summary features. We have repeatedly argued that AND-like feature combinations are essential for our understanding of the visual system [27,30,34,35,36,28]. The increased selectivity of nonlinear AND operators, as compared to their linear counterparts, is a prerequisite for the usefulness of integrals over the respective responses [30,28]. We have shown that such integrals of AND features are relevant for the understanding of texture perception [37], of numerosity estimation [30], and of invariance in general [28]. Recently, integrals over AND-like feature combinations in form of auto- and cross-correlation functions have been suggested for the understanding of peripheral vision [4,16,17].

A somewhat surprising point is that linear summation and cortical gain control, two widely accepted properties of cortical neurons, are the only requirements for the computation of ensemble properties. These functions are already available at early stages of the cortex, but also in other cortical areas [23]. The computation of ensemble properties may thus be an ubiquitous phenomenon in the cortex.

## Acknowledgement

This work was supported by DFG, SFB/TR8 Spatial Cognition, project A5-[ActionSpace].

## References

- [1] S. C. Dakin and R. J. Watt. The computation of orientation statistics from visual texture. *Vision Res*, 37(22):3181–3192, 1997.
- [2] D. Ariely. Seeing Sets: Representation by Statistical Properties. *Psychol Sci*, 12(2):157–162, 2001.
- [3] Lin Chen. The topological approach to perceptual organization. *Visual Cognition*, 12(4):553–637, 2005.
- [4] B. Balas, L. Nakano, and R. Rosenholtz. A summary-statistic representation in peripheral vision explains visual crowding. *J Vis*, 9(12):13.1–18, 2009.
- [5] G. A. Alvarez. Representing multiple objects as an ensemble enhances visual cognition. *Trends Cog Sci*, 15(3):122–31, 2011.
- [6] R. Rosenholtz, J. Huang, and K. Ehinger. Rethinking the role of top-down attention in vision: effects attributable to a lossy representation in peripheral vision. *Front Psychol*, 3:13, 2012.
- [7] S. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381(6582):520–522, 1996.
- [8] A. Oliva and A. Torralba. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *International Journal of Computer Vision*, 42(3):145–175, 2001.
- [9] E. M. Brannon. The representation of numerical magnitude. *Curr Opin Neurobiol*, 16(2):222–9, 2006.
- [10] J. Hegde and D.J. Felleman. Reappraising the Functional Implications of the Primate Visual Anatomical Hierarchy. *The Neuroscientist*, 13(5):416–421, 2007.
- [11] D. Navon. Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9(3):353–383, 1977.
- [12] M. R. Greene and A. Oliva. The Briefest of Glances. *Psychol Sci*, 20(4):464–472, 2009.

- [13] J. Hegdé. Time course of visual perception: coarse-to-fine processing and beyond. *Prog Neurobiol*, 84(4):405–39, 2008.
- [14] M. Fabre-Thorpe. The characteristics and limits of rapid visual categorization. *Front Psychol*, 2:243, 2011.
- [15] M. J-M Macé, O. R. Joubert, J-L. Nespoulous, and M. Fabre-Thorpe. The time-course of visual categorizations: you spot the animal faster than the bird. *PloS one*, 4(6):e5927, 2009.
- [16] J. Freeman and E. P. Simoncelli. Metamers of the ventral stream. *Nature neuroscience*, 14(9):1195–1201, 2011.
- [17] H. Strasburger, I. Rentschler, and M. Jüttner. Peripheral vision and pattern recognition: a review. *J Vis*, 11(5):13, 2011.
- [18] A. Nieder, D. J. Freedman, and E. K. Miller. Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, 297(5587):1708–11, 2002.
- [19] H. J. Gross, M. Pahl, A. Si, H. Zhu, J. Tautz, and S. Zhang. Number-based visual generalisation in the honeybee. *PloS one*, 4(1):e4263, 2009.
- [20] J. D. Roitman, E. M. Brannon, and M. L. Platt. Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS biology*, 5(8):e208, 2007.
- [21] J. Ross and D. C. Burr. Vision senses number directly. *Journal of vision*, 10(2):10.1–8, 2010.
- [22] D. G. Albrecht and D. B. Hamilton. Striate cortex of monkey and cat: contrast response function. *J Neurophysiol*, 48(1):217–237, Jul 1982.
- [23] M. Carandini and D. J. Heeger. Normalization as a canonical neural computation. *Nature Reviews Neurosci*, 13:51–62, Jul 2012.
- [24] R. Rosenholtz, J. Huang, A. Raj, B. J. Balas, and L. Ilie. A summary statistic representation in peripheral vision explains visual search. *J Vis*, 12(4):1–17, 2012.
- [25] H.L. Resnikoff and R.O. Wells. *Mathematics in Civilization*. Popular Science Series. Dover, 1984.
- [26] E. H. Adelson and J. R. Bergen. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A*, 2(2):284–99, 1985.
- [27] C. Zetzsche and E. Barth. Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Res*, 30(7):1111–1117, 1990.
- [28] C. Zetzsche and U. Nuding. Nonlinear and higher-order approaches to the encoding of natural scenes. *Network*, 16(2–3):191–221, 2005.
- [29] D.J. Heeger. Normalization of cell responses in cat striate cortex. *Visual Neurosci*, 9(2):181–198, 1992.
- [30] C. Zetzsche and E. Barth. Image surface predicates and the neural encoding of two-dimensional signal variations. In B. E. Rogowitz and Jan P. A., editors, *Proc SPIE*, volume 1249, pages 160–177, 1990.
- [31] S. Dehaene and J. P. Changeux. Development of elementary numerical abilities: a neuronal model. *J. Cogn. Neurosci.*, 5(4):390–407, 1993.
- [32] I. Stoianov and M. Zorzi. Emergence of a ‘visual number sense’ in hierarchical generative models. *Nat Neurosci*, 15(2):194–6, 2012.
- [33] M. Ferraro E. Barth and C. Zetzsche. Global topological properties of images derived from local curvature features. In L. P. Cordella C. Arcelli and G. Sanniti di Baja, editors, *Visual Form 2001. Lecture Notes in Computer Science*, pages 285–294, 2001.
- [34] C. Zetzsche, E. Barth, and B. Wegmann. The importance of intrinsically two-dimensional image features in biological vision and picture coding. In A. B. Watson, editor, *Digital images and human vision*, pages 109–138. MIT Press, Cambridge, MA, 1993.
- [35] G. Krieger and C. Zetzsche. Nonlinear image operators for the evaluation of local intrinsic dimensionality. *IEEE Transactions Image Processing*, 5:1026–1042, 1996.
- [36] C. Zetzsche and G. Krieger. Nonlinear mechanisms and higher-order statistics in biological vision and electronic image processing: review and perspectives. *J Electronic Imaging*, 10(1):56–99, 2001.
- [37] E. Barth, C. Zetzsche, and I. Rentschler. Intrinsic 2D features as textons. *J. Opt. Soc. Am. A*, 15(7):1723–1732, 1998.

